

TEKNILLINEN KORKEAKOULU

Sähkö- ja tietoliikennetekniikan osasto

Mikko Ojamo

**Käytettävyydestien korvattavuus heuristisella
arvioinnilla sisällönhallintajärjestelmissä**

**Diplomityö, joka on jätetty opinnäytteenä tarkastettavaksi diplomi-insinöörin
tutkintoa varten Espoossa 27.5.2005**

Tekijä:	Mikko Ojamo		
Työn nimi:	Käytettävyystestien korvattavuus	heuristisella	arvioinnilla sisällönhallintajärjestelmissä
Päivämäärä:	27.5.2005	Sivuja: 72 + liitteet	
Osasto:	Sähkö- ja tietoliikennetekniikan osasto	Professuuri: T-111	
Työn valvoja:	Petri Vuorimaa		
Työn ohjaaja:	Pirkka Paronen		
<p>Tässä diplomityössä pyrittiin selvittämään voidaanko heuristisella arvioinnilla korvata käytettävyystestit sisällönhallintajärjestelmän käytettävyyttä arvioitaessa. Käytettävyystesteissä tutkittavaa järjestelmää käyttävät useat testikäyttäjät ja menetelmä vaatii runsaasti aikaa ja resursseja. Heuristisessa arvioinnissa muutama käytettävyysasiantuntija pyrkii arvioimaan mihin ongelmiin todelliset käyttäjät järjestelmää käyttäessään mahdollisesti törmäisivät. Kun arvoinnit voidaan suorittaa ja analysoida jopa yhdessä päivässä, tarjoaa heuristinen arviointi halvemman keinon käytettävyysongelmien kartoittamiselle. Epäselvää oli kärsiikö käytettävyuden arvioinnin laatu kustannussäästöjen mukana eli jääkö heuristisella arvioinnilla löytymättä sellaisia käytettävyysongelmia, joita käytettävyystesteillä huomataan.</p> <p>Tutkimuksessa suoritettiin arvoinnit molemmilla menetelmillä yhteneväsillä testijärjestelyillä, jolloin tuloksia voitiin luotettavasti vertailla keskenään. Paljastui, että heuristisella arvioinnilla löydettiin lähes kaikki käytettävyysongelmat, mitkä löydettiin käytettävyystesteilläkin. Lisäksi heuristisessa arvioinnissa löytyi paljon sellaisia ongelmia, joita ei käytettävyystesteillä löydetty. Näiden tulosten perusteella voidaan sanoa, että heuristista arviointia voidaan käyttää käytettävyystestien sijaan.</p>			
Avainsanat:	Käytettävyys,	heuristinen	arviointi, käytettävyystesti, sisällönhallintajärjestelmä

Author:	Mikko Ojamo		
Name of the thesis:	Heuristic analysis as a substitute method for usability tests when evaluating content management systems		
Date:	27.5.2005	Number of pages:	72 + appendices
Department:	Department of Electrical and Communications Engineering		
Professorship:	T-111		
Supervisor:	Petri Vuorimaa		
Instructor:	Pirkka Paronen		
<p>In this thesis, the goal was to find out, whether the heuristic analysis could be used as substitute method for usability tests when evaluating content management system. In usability tests, the method requires many test users, and therefore lots of time and resources. In heuristic analysis the usability experts try evaluate what kind of problems actual users could face when using the system. These evaluations can be done and analyzed within a day, and therefore heuristic analysis offers less expensive method for usability evaluation. The question was does the quality of research suffer with the cost savings. Are all the problems that are found with usability test found also with heuristic analysis.</p> <p>In this research, the usability evaluations were done with both methods in identical test settings. Therefore, the test results were comparable. It was discovered that heuristic evaluation found all the usability problems found with usability tests and also many others that weren't found with usability tests. These results show that heuristic evaluations can be used as substitute method for usability tests.</p>			
Keywords: Usability, heuristic analysis, usability test, content management system			

Sisältö

1	JOHDANTO.....	1
1.1	TAUSTAA.....	1
1.2	TUTKIMUSTAVOITTEET.....	2
1.3	TYÖN RAKENNE	3
2	KÄYTETTÄVYYS JA KÄYTETTÄVYYDEN ARVIOINTI	4
2.1	HEURISTINEN ARVIOINTI.....	5
2.1.1	<i>Heuristiikka</i>	<i>6</i>
2.1.2	<i>Arvioinnin toteutus</i>	<i>8</i>
2.1.3	<i>Käytettävyyssongelmien luokittelu.....</i>	<i>9</i>
2.1.4	<i>Menetelmän vahvuudet ja heikkoudet.....</i>	<i>11</i>
2.2	KÄYTETTÄVYYSTESTIT.....	13
2.2.1	<i>Testien rakenne.....</i>	<i>14</i>
2.2.2	<i>Löydettyjen käytettävyyssongelmien raportointi.....</i>	<i>16</i>
2.2.3	<i>Käytettävyyssongelmien luokittelu.....</i>	<i>17</i>
2.2.4	<i>Menetelmän vahvuuksia ja heikkouksia</i>	<i>18</i>
2.3	KÄYTETTÄVYYDENARVIOINTIMENETELMIEN VERTAILUN TEORIA	20
2.3.1	<i>Löydettyjen käytettävyyssongelmien määrän vertailu.....</i>	<i>21</i>
2.3.2	<i>Arviointimenetelmien vertailun haasteita.....</i>	<i>22</i>
2.3.3	<i>Kustannustehokkuus</i>	<i>23</i>
3	SISÄLLÖNHALLINTAJÄRJESTELMÄT.....	25
3.1	RAKENNE JA KÄYTTÖ	26
3.2	SISÄLLÖNHALLINTAJÄRJESTELMÄN ETUJA	27
3.3	SISÄLLÖNHALLINTAJÄRJESTELMIEN KÄYTETTÄVYYDEN ARVIOINTI.....	28
3.3.1	<i>Järjestelmään lukittautumisen vaikutukset.....</i>	<i>29</i>
3.4	NAS-SISÄLLÖNHALLINTAJÄRJESTELMÄ	30
3.4.1	<i>Järjestelmän rakenne.....</i>	<i>30</i>

3.4.2	<i>Järjestelmän käyttö</i>	32
3.4.3	<i>Järjestelmän ylläpitäjät</i>	34
4	TESTIEN RAKENNE JA TOTEUTUS	36
4.1	TESTISIVUSTO.....	36
4.2	TESTITEHTÄVIEN SUUNNITTELU	37
4.3	HEURISTISEN ARVIOINNIN SUUNNITTELU JA TOTEUTUS	39
4.4	KÄYTETTÄVYYSTESTIEN SUUNNITTELU JA TOTEUTUS	40
4.4.1	<i>Testikäyttäjien valinta</i>	40
4.4.2	<i>Testin rakenne</i>	41
4.4.3	<i>Testitilanne</i>	41
4.4.4	<i>Pilottitesti</i>	42
5	TULOKSET	43
5.1	HEURISTISEN ARVIOINNIN TULOKSET	43
5.1.1	<i>Löydetyt käytettävyysoingelmat</i>	45
5.2	KÄYTETTÄVYYSTESTIEN TULOKSET	47
5.2.1	<i>Löydetyt käytettävyysoingelmat</i>	51
5.3	ARVIOINTIMENETELMIEN TULOSTEN VERTAILU.....	53
5.3.1	<i>Tulosten vertailu tehtävittäin</i>	53
5.3.2	<i>Vertailun yhteenveto</i>	54
6	ANALYYSI	59
6.1	TULOSTEN VASTAAVUUS AIEMPIIN TUTKIMUKSIIN	59
6.1.1	<i>Virheelliset hälytykset</i>	61
6.1.2	<i>Arvioinnin puutteet</i>	61
6.1.3	<i>Testien puutteet</i>	62
7	JOHTOPÄÄTÖKSET JA YHTEENVETO	64
7.1	VASTAUS TUTKIMUSKYSYMYKSEEN	64
7.2	JATKOTUTKIMUKSEN AIHEITA	65

8	LÄHTEET.....	66
9	LIITTEET.....	73

Lyhenteitä ja määritelmiä

Alennettu käytettävyyden tutkiminen

Discount usability engineering. Jacob Nielsenin kehittämä käytettävyyden arviointimenetelmä, jossa ei pyritä täydellisyyteen vaan löytämään edes osa käytettävyysongelmista. Yhdistää käyttäjien ja tehtävien tarkkailun, skenariot, yksinkertaistetut käytettävyyshetimit ääneen ajattelulla ja heuristisen arvioinnin.

Extranet

Extranet on intranet, jonka rajoja on laajennettu yrityksen ulkopuolelle. Se pohjautuu internet-tekniikkaan, mutta käyttöoikeudet on rajattu vain yrityksen henkilökunnalle sekä esimerkiksi yhteistyökumppaneille ja asiakkaille.

Intranet

Intranet on julkisesta verkosta esimerkiksi palomureilla eristetty yrityksen sisäinen verkko, johon on pääsy vain yrityksen henkilökunnalla

Kvalitatiivinen aineisto

Pieni määrä aineistoa, josta tehdään tulkintoja. Pohjautuu usein haastatteluihin tai havainnointeihin.

Kvantitatiivinen aineisto

Pohjautuu suureen aineistomäärään, joka on kerätty esimerkiksi tarkoin

	mittauksin. Johtopäätökset voidaan rakentaa tilastotieteen menetelmin.
Sisällönhallintajärjestelmä	WWW-sivujen ylläpitoon tarkoitettu järjestelmä, jota käytetään esimerkiksi internet-selaimen avulla.
Sovellus	Ohjelma tai joukko ohjelmia, jotka prosessoivat dataa käyttäjälle. Esimerkkinä tekstinkäsittely- tai taulukkolaskentaohjelmat.
Virheellinen hälytys	Käytettävyyssongelmaksi luokitellaan sellainen toiminto tai käyttöliittymän osanen, joka ei aiheuta todellisessa käyttötilanteessa käytettävyyssongelmia.
Verkko-ohjelmisto	Ohjelmat, joita ei ole asennettu käyttäjän koneelle vaan käytetään esimerkiksi internet-selaimen avulla verkkoyhteyden välityksellä.
Villikortti-ilmiö	Wildcard effect. Gray ja Salzmanin [Gray & Salzman 1998] raporttoima ilmiö, jossa käytettävyystesteissä pienellä testikäyttäjämäärällä yhden testikäyttäjän huomattavasti keskimääräistä parempi tai huonompi suoritus vaikuttaa tulosten luotettavuuteen.

1 Johdanto

1.1 Taustaa

Käytettävyyden testaus on tärkeä osa ohjelmistotuotteiden kehitystyötä. Huolimaton suunnittelu voi johtaa järjestelmiin, jotka ovat vaikeita oppia ja käyttää, luovat ylimääräistä työtaakkaa käyttäjille ja pahimmillaan altistavat vakaville virhetilanteille [Roth et al. 2002]. Käytettävyyden testausta voidaan pitää osana laadunvarmistusta, sillä siinä pyritään varmistamaan, että tuote on helppo oppia, se on hyödyllinen sekä mukava että helppo käyttää [Gould & Lewis 1983]. Käytettävyyden testaus on usein yksi merkittävimmistä osista koko tuotteen elinkaaren kattavassa käytettävyyden huomioinnissa. Testit voivat olla myös merkittävä kustannuserä ja vaatia huomattavasti aikaa, varsinkin jos ne halutaan suorittaa todellisilla käyttäjillä. Tähän ei kaikissa projekteissa ole mahdollisuuksia, joten tarve kevyempiin menetelmiin on ilmeinen.

Kustannus- ja aikataulusäästöjen hakeminen käytettävyyden testauksesta on kuitenkin riskialtista toimintaa. On muistettava, että jokaisen tuotteen käytettävyyks tulee arvioitua ennemmin tai myöhemmin. Jos tuotteen valmistaja ei tee huolellista arviointia ennen julkaisua, käytettävyyssongelmat tulevat esiin vasta asiakkaiden käyttäessä tuotetta. Julkaisun jälkeen muutosten tekeminen on yleensä huomattavasti vaikeampaa kuin jos virheet huomataan heti projektin alussa. Välttämättä jälkikäteen kaikkia virheitä ei voida korjata ollenkaan. Lisäksi on huomioitava käytettävyyssongelmista aiheutuva imagohaitta tuotteen valmistajalle sekä mahdollinen koulutus- ja käyttötukikulujen kasvu. [Nielsen 1993, Rosenberg 2004]

Ennen tietyn testausmenetelmän käyttöönottoa on siis tärkeää selvittää soveltuuko testityyppi käyttötarkoitukseen ja voidaanko sillä saavuttaa kustannus- ja aikataulusäästöjä niin, että tulokset säilyvät kuitenkin vaaditulla tasolla.

Tämä diplomityö sai alkunsa Systems Garden Oy:n tarpeesta tutkia heidän tuotteensa, NAS-sisällönhallintajärjestelmän, käytettävyyttä. Pienenä yrityksenä, henkilökunnan kokonaismäärä 8 henkilöä lokakuussa 2004, Systems Gardenilla ei ole mahdollisuutta itse toteuttaa käytettävyystestejä kaikissa tuotekehitysprojekteissa, sillä ne vaativat vähintään yhden työntekijän keskittyvän testityöhön useaksi viikoksi. Oli siis selkeä tarve hakea kevyempiä ratkaisuja.

Heuristinen arviointi valikoitui vaihtoehdoksi käytettävyystesteille aikaisempia tutkimuksia läpikäydessä [mm. John & Marks 1997, Jeffries et al. 1991, Law & Hvannberg 2002]. Oli kuitenkin selvittävä miten menetelmät soveltuvat sisällönhallintajärjestelmien tutkimiseen, sillä aikaisemmissa tutkimuksissa sisällönhallintajärjestelmien erityispiirteitä ei oltu käsitelty ollenkaan. Sisällönhallintajärjestelmiin ja NAS:iin palataan tarkemmin luvussa 3.

1.2 Tutkimustavoitteet

Edellisessä luvussa esitettyjen lähtökohtien pohjalta tutkimustavoite voitiin asettaa seuraavasti:

Voidaanko käytettävyystestaus korvata heuristisella arvioinnilla sisällönhallintajärjestelmän käytettävyyttä arvioitaessa?

Jotta tähän kysymykseen voidaan vastata tutkimuksen tavoitteiksi asetettiin:

- Tutkia miten käytettävyystestejä ja heuristista arviointia voidaan käyttää luotettavasti keskenään sisällönhallintajärjestelmän käytettävyyttä tutkittaessa (luvut 2.3 ja 3.3)
- Suunnitella ja toteuttaa käytettävyyden arviointi tutkittavalle järjestelmälle edellä mainittuja menetelmiä käyttäen (luku 4)
- Verrata tuloksien perusteella arviointimenetelmiä (luku 5)

Esiteltäviä arviointimenetelmiä voidaan käyttää monenlaisten järjestelmien testaukseen. Tässä työssä aihetta tarkastellaan puhtaasti verkko-ohjelmiston ja tarkemmin sanottuna sisällönhallintajärjestelmän tarpeiden kautta. Näin ollen tämän työn tulokset eivät välttämättä ole sovellettavissa muihin ohjelmistotyyppeihin. Sisällönhallintajärjestelmien käyttäjistä tässä työssä keskitytään sisällön ylläpitäjän ja pääkäyttäjän näkökulmiin. Käytettävyyden arvioinnit pohjustetaan näiden käyttäjien tarpeet erityisesti huomioiden. Tarkemmin nämä käyttäjäroolit esitellään luvussa 3.4.3.

Arviointien tuloksena syntyy myös käsitys testattavan järjestelmän kehityskohteista. Tämän diplomityön osalta ne ovat kuitenkin sivutuotteita, sillä painopiste on menetelmien vertailussa. Käytettävyyssongelmien korjaukset jäävät näin ollen työn ulkopuolelle. Korjausehdotuksiin työni ottaa kantaa, mutta tässäkin näkökulma on lähinnä miten verrattavat menetelmät tukevat korjausehdotusten tekoa.

1.3 Työn rakenne

Työn toisessa ja kolmannessa luvussa pyritään esittelemään aiheeseen liittyvää teoriaa aikaisempia tutkimuksia läpikäyden. Esitellään tarkemmin tutkimuksessa käytetyt menetelmät: heuristisen arvioinnin ja käytettävyyssitestit. Toisessa luvussa käydään myös läpi menetelmien vertailuun liittyviä haasteista. Kolmannessa luvussa syvennytään sisällönhallintajärjestelmiin sekä esitellään tässä tutkimuksessa käytetty NAS-sisällönhallintajärjestelmä.

Neljäs luku esittelee työn käytännön osuuden. Tavoitteena oli suunnitella yhtenäinen testirakenne ja suorittaa se sekä heuristisella arvioinnilla että käytettävyyss testeillä. Viidennessä luvussa käydään tulokset, joita analysoidaan kuudennessa luvussa. Lopuksi seitsemännessä luvussa tehdään yhteenveto ja johtopäätökset.

2 Käytettävyys ja käytettävyyden arviointi

Käytettävyyden arviointimenetelmät voidaan jaotella Taulukko 1:ssä osoitetulla 2x2 luokittelujärjestelmällä [Connell 2000]. Luokittelu on muunnelmä Whitefieldin tutkimusryhmän esittämästä mallista [Whitefield et al. 1991].

Taulukko 1. Käytettävyysmenetelmien 2x2 luokittelu

	Ilman käyttäjää	Käyttäjän kanssa
Järjestelmän kanssa	Arviointimenetelmät	Havainnointimenetelmät
Ilman järjestelmää	Analyytiset menetelmät	Kyselymenetelmät

Arviointimenetelmiä käytettäessä käytettävyysarvioinnit toteutetaan toimivalla järjestelmällä, prototyypillä tai määrittelyksien avulla, mutta ilman todellisia käyttäjiä. Arvioinnin suorittavat käytettävyysasiantuntijat tai muut henkilökunnan jäsenet, jotka käyttävät apunaan jonkinlaisia arviointikriteereitä. Menetelmissä asetutaan usein käyttäjän asemaan ja pyritään arvioimaan mihin ongelmiin todellinen käyttäjä saattaisi törmätä. Heuristinen arviointi on yksi arviointimenetelmistä.

Havainnointimenetelmissä tarkkaillaan käyttäjän ja järjestelmän vuorovaikutusta. Paikka voi vaihdella käytettävyyslaboratoriosta käyttäjän työpisteeseen. Tilanteesta voidaan kerätä tietoa järjestelmän ja käyttäjän suorituksesta ja lisäksi käyttäjää saatetaan pyytää selostamaan ääneen toimintaansa ja ajatuksiaan. Mittausten lisäksi käyttäjää voidaan haastatella vapaamuotoisesti sekä ennen testejä että jälkeen. Tässä työssä esitelty käytettävyystesti kuuluu havainnointimenetelmiin.

Analyttisissä menetelmissä asiantuntijat tai suunnitteluryhmän jäsenet ennustavat järjestelmän käyttöä erilaisten mallien tai määritysten avulla, jotka on voitu tehdä järjestelmästä, käyttöliittymästä tai käyttäjistä. Tavoitteena on tehdä ennusteita esimerkiksi virhetilanteista ja käyttäjien tarpeista. Ennusteissa ei käytetä apua oikeita käyttäjiä eikä tutkita toimivaa järjestelmää. Esimerkki analyttisestä menetelmästä on GOMS (Goals, Operations, Methods and Selection rules) [John 1995].

Kyselymenetelmissä käyttäjiltä kerätään erilaisten kyselylomakkeiden, yksilö- ja ryhmähaastatteluiden avulla kokemuksia järjestelmän käytöstä. Kerätty tieto voi sisältää mielipiteitä järjestelmästä ja kokemuksia ongelmatilanteista

Tässä työssä tarkasteluun rajattiin vain arviointimenetelmiin kuuluvaan käytettävyyystesteihin ja havainnointimenetelmiin kuuluvaan heuristiseen arviointiin. Syynä valintaan oli se, että analyttiset menetelmät soveltuvat paremmin tilanteeseen, jossa ei ole valmiista tuotetta. Tässä tutkimuksessa tällaista tilannetta ei ollut. Kyselymenetelmissä kerätty aineisto taas on epäluotettavampaa kuin havainnointimenetelmissä saatu aineisto. Kyselyissä käyttäjät eivät aina anna todenmukaisia vastauksia ja lisäksi vastaukset eivät anna suoraa tietoa käyttöliittymästä vaan kertovat lähinnä käyttäjän mielipiteestä [Nielsen 1993]. On eri asia tietää mitä käyttäjä tekee, kuin mitä hän luulee tekevänsä ja tämä saadaan parhaiten selville havainnointimenetelmillä kuten käytettävyystesteillä.

2.1 Heuristinen arviointi

Heuristinen arviointi (heuristic evaluation) on vapaamuotoinen arviointimenetelmä, jossa käytettävyyso ongelmia pyritään löytämään tutkimalla käyttöliittymää ja vertaamalla sitä yleiseen suunnittelusäännöstöön eli heuristiikkaan [Nielsen 1995]. Se on yksi ilman todellisia käyttäjiä toteutettavista arviointimenetelmistä. Muita tähän ryhmään kuuluvia metodeja ovat esimerkiksi kognitiivinen läpikäynti (cognitive walkthrough) [Lewis et al. 1990] ja väitteiden analyysi (claims analysis) [Carroll & Rossom 1992].

Käyttöliittymää suunniteltaessa suunnittelija tulee luontevasti arvioineeksi mitkä elementit ovat hyviä tai huonoja käytettävyyden kannalta. Usein nämä arviot pohjautuvat suunnittelijan kokemukseen tai tuntemuksiin, mutta apuna on voitu käyttää myös yksityiskohtaisia suunnitteluohjeita. Esimerkiksi Microsoft on julkaissut erillisen satoja sivuja paksun suunnitteluoppaan Windows-ohjelmien käyttöliittymän suunnitteluun [Microsoft Corporation 1999]. Tällaisten massiivisten oppaiden käyttö suunnittelun apuna on varsin vaivalloista, joten Nielsenin ja Molichin lähtökohta heuristista arviointia kehittäessään oli käyttää arvioinnissa vain hyvin yleisiä suunnittelusääntöjä ja karsia nämä säännöt minimiin [Molich & Nielsen 1990]. Heidän kehittämänsä kymmenen kohdan suunnittelusäännöstöä käsitellään tarkemmin kohdassa 2.1.1.

Heuristisen arvioinnin suorittamiseen ei välttämättä tarvita edes toimivaa sovellusta vaan menetelmää voidaan hyödyntää jo esimerkiksi käyttöliittymähahmotelmilla tai paperiprototyypeillä [Nielsen 1993, 159]. Näin käytettävyysongelmia on mahdollista karsia heti suunnittelun alkuvaiheessa, jolloin myös muutostkustannukset pysyvät kurissa. Menetelmän helpon läpikäynnin ansiosta arviointi voidaan toistaa myös useaan kertaan suunnitteluprojektin eri vaiheissa.

2.1.1 Heuristiikka

Heuristisessa arvioinnissa käytettävä heuristiikka tarkoittaa kohtuullisen pientä määrää yleisluontoisia käytettävyysperiaatteita [Nielsen 1994]. Yksi yleisimmin käytettyjä on Nielsenin ja Molichin [Nielsen 1990, s.20] kehittämä kymmenen kohdan heuristiikka, joka on suunniteltu erityisesti tietokoneohjelmien käyttöliittymän arviointiin.

1. **Yksinkertainen ja luonnollinen vuoropuhelu.** Vuoropuhelut eivät saisi sisältää tarpeetonta tai harvoin tarvittavaa tietoa. Jokainen ylimääräinen tieto vuoropuhelussa kilpailee tärkeiden tietojen kanssa näkyvyydestä ja heikentää

niiden suhteellista näkyvyyttä. Kaikki tieto pitäisi näyttää luonnollisessa ja loogisessa järjestyksessä.

2. **Puhu käyttäjän kieltä.** Vuoropuhelu tulisi esittää käyttäen selkeitä ja käyttäjälle tuttuja sanoja, sanontoja ja konsepteja eikä järjestelmäkeskeisiä termejä.
3. **Minimoi käyttäjän muistikuorma.** Käyttäjä ei saisi joutua muistamaan tietoa yhdestä vuoropuhelun kohdasta toiseen. Järjestelmän käyttöohjeiden tulisi olla näkyvillä tai helposti saatavilla, kun niitä tarvitaan.
4. **Yhtenäisyys.** Käyttäjä ei saisi joutua miettimään tarkoittavatko eri sanat, tilanteet tai toiminnot samaa asiaa.
5. **Palaute.** Järjestelmän tulisi aina tiedottaa käyttäjälle mitä on tapahtumassa riittävällä palautteella riittävän nopeasti.
6. **Selvästi merkityt poistumistiet.** Käyttäjät valitsevat usein järjestelmän toimintoja vahingossa ja tarvitsevat selkeästi merkittyjä ”häätäpoistumisteitä” poistuaan ei-toivotuista toiminnoista ilman pitkiä vuoropuheluita.
7. **Oikotiet.** Kiihdyttimet, joita noviisikäyttäjät eivät huomaa, voivat nopeuttaa asiantuntijakäyttäjien työtä niin, että järjestelmä pystyy palvelemaan sekä aloittelevia että edistyneitä käyttäjiä.
8. **Hyvät virheilmoitukset.** Virheilmoitukset tulisi esittää selkeästi sanoin (ei koodeja) ja niiden tulisi määrittää selkeästi ongelma sekä ehdottaa siihen ratkaisua.

9. **Estä virheet.** Hyviä virheilmoituksia tärkeämpää on huolellinen suunnittelu, joka estää virheiden syntymisen.
10. **Aputoiminto ja dokumentaatio.** Vaikka on parempi, jos järjestelmää voi käyttää ilman dokumentaatiota, voi olla tarpeen tarjota aputoiminto ja dokumentaatio. Tällaisen tiedon tulisi olla helposti etsittävissä, keskittynyt käyttäjän tehtävään, luetella konkreettiset suoritettavat vaiheet, eikä se saisi olla liian iso.

2.1.2 Arvioinnin toteutus

On suositeltavaa, että heuristisen arvioinnin suorittavat käytettävyysasiantuntijat, mutta tämä ei ole välttämätöntä. Tutkimuksissa parhaimpiin tuloksiin päästiin, kun arvioijilla oli asiantuntemusta sekä käytettävyydestä että arvioitavasta järjestelmästä. Novisiarvioijatkin pääsevät varsin hyviin tuloksiin, mutta arvioijia tarvitaan useampia. Asiantuntijoiden suorittamissa arvioinneissa arviointi on tehokkainta, kun osallistujia on kolmesta viiteen. Tätä suuremmilla määrillä saavutettujen tulosten suhde testin aiheuttamiin kustannuksiin alkaa laskea selvästi, sillä testaajat löytävät lähinnä ongelmakohtia, jotka joku toinen testaajista on jo havainnut. [Nielsen 1992]

Varsinaisessa käyttöliittymän arvioinnissa arvioijat käyvät ensin itsenäisesti käyttöliittymän läpi vähintään kahteen kertaan [Nielsen 1993, 158]. Ensimmäisellä kerralla käyttöliittymää käydään läpi yleisellä tasolla selaillen ja tutkien ja toisella kerralla jokainen käyttöliittymäelementti tutkitaan yksityiskohtaisesti. Läpikäynneissä voidaan käyttää myös valmiita tehtäviä, joita arvioijien on pyrittävä sovelluksella suorittamaan [Karat et al. 1992]. Itsenäisen arviointiin käytetään kerrallaan tunnista kahteen. Kaikki löydöt kirjataan paperille ja lopuksi arvioijat kokoontuvat yhdessä tutkimaan tuloksia. Yhteenvetopalaveri on tärkeä myös siinä suhteessa, että arvioijat voivat vapaasti keskustella järjestelmän käytöstä ja antaa myös positiivista palautetta. Menetelmä ei tähän muuten kannusta, joten positiiviset seikat voivat helposti jäädä mainitsematta. Käytettävyysongelmista laaditaan raportti,

jossa kaikki ongelmat on eritelty ja kohdistettu mitä heuristiikan kohtaa tai kohtia se rikkoo. Loppuraportissa voidaan myös ehdottaa korjauskeinoja.

2.1.3 Käytettävyysongelmien luokittelu

Arvioinnin tuloksena syntyy lista löydettyistä käytettävyysongelmista, jotka vaihtelevat hyvinkin pienistä yksityiskohdista suurempiin useita käytettävyysperiaatteita rikkoviin virheisiin. Ongelmat on hyvä luokitella käytettävyysongelman vakavuuden suhteen, jotta jatkotoimista päättäminen on helpompaa. Vakavimpien ongelmakohtien nostaminen esille on tärkeää sillä, jos korjaustoimista vastaaville toimitetaan vain suunnaton lista ongelmakohdista, joista suurin osa on hyvin mitättömiltä vaikuttavia vikoja, tämä saattaa nostattaa hylkäysreaktion ja kaikki uudistukset jäädä tekemättä [Molich et al. 1999]. Luokittelun avulla voidaan paremmin arvioida saadaanko korjauksesta niin merkittävää hyötyä verrattuna muutoksesta aiheutuviin kustannuksiin, että järjestelmään tarvittavat muutokset ovat perusteltuja.

Nielsenin [Nielsen 1994b] mukaan käytettävyysongelman vakavuus määräytyy kolmesta osa-alueesta: toistuvuus (kuinka monta kertaa ongelma ilmenee), vaikutus (onko käyttäjän helppo vai vaikea selvittää ongelmasta), pysyvyys (onko ongelma kertaluontoinen, joka myöhemmin on helppo ohittaa kun tietää ratkaisun vai juuttuuko käyttäjä siihen toistuvasti). Lisäkohtana Nielsen mainitsee markkinavaikutuksen, sillä laajempi huomio voi nostaa paineita käytettävyysongelman korjaukseen, vaikka edellisten kohtien avulla arvioituna ongelman vakavuus ei olisi kovin korkea.

Jokaisen arvioijan tulisi luokitella löytämänsä käytettävyysongelmat itsenäisesti arvioinnin jälkeen. Luokittelua helpottamaan voidaan käyttää esimerkiksi valmista kyselylomaketta. Yksittäisen arvioijan luokittelu ei kuitenkaan ole kovin luotettava, joten lopulliseen raporttiin tulee tulokset koostaa usean arvioijan luokitteluista. Vakavammat käytettävyysongelmat tulevan näin paremmin esille. [Nielsen 1994b] Nielsen esittää myös, että lopullinen luokitteluasteikko koostetaan yksittäisten

arviointiasteikkojen keskiarvosta, jolloin erot eri arvioijien asteikoissa tasoittuvat [Nielsen 1993].

Nielsenin esittelemä [Nielsen 1993, 103] yksinkertainen luokitteluasteikko koostuu viidestä luokasta.

1. ei käytettävyysongelma
2. kosmeettinen käytettävyysongelma. Ei tarvitse korjata ellei projektissa ole ylimääräistä aikaa
3. vähäinen käytettävyysongelma. Korjaukselle tulee antaa alhainen prioriteetti
4. merkittävä käytettävyysongelma. Tärkeää korjata korkealla prioriteetilla
5. käytettävyyskatastrofi. Ehdotonta korjata ennen julkaisua

Karat tutkimuskumppaneineen [Karat et al. 1992] esittelee yhdeksi mahdolliseksi luokittelujärjestelmäksi PSC-järjestelmän (Problem Severity Classification) eli ongelmien vakavuusluokittelun. Siinä käytettävyysongelmia käsitellään kahden asteikon avulla. Yhdellä akselilla on ongelman vaikutus käyttäjän mahdollisuuteen suorittaa tehtävä ja toisella akselilla on toistuvuus (kuinka suuri prosenttiosuus arvioijista törmäsi ongelmaan). Vaikutus-akseli oli jaettu kolmeen arvoon (korkea, kohtalainen, alhainen) ja toistuvuus-akseli kahteen arvoon (korkea, kohtalainen). Nämä yhdistämällä saadaan PSC-luokitus, jonka skaala on yhdestä kolmeen. Asteikossa 1 tarkoittaa vakavaa käytettävyysongelmaa. Vaikutus-akselilla korkea arvo asetetaan ongelmalle, joka estää tehtävän suorituksen, kohtalainen merkitsee suurta haittaa tehtävän suoritukseen ja alhainen pientä ongelmaa tai tehottomuutta tehtävän suorittamisessa. Karatin tutkimuksessa toistuvuus-asteikolla kohtalaiseksi määäräytyivät ongelmat, joihin törmäsi kaksi käyttäjää (33%) ja korkeaksi ongelmat, joihin törmäsi kolme tai useampi käyttäjä (50%). Lopulliseksi PSC-luokituksen ongelmat saavat Taulukko 2:ssa osoitetulla jaolla.

Taulukko 2. PSC-luokitus

Vaikutus tehtävään	Toistuvuus	
	Korkea	Kohtalainen
Korkea	1	1
Kohtalainen	1	2
Matala	2	3

2.1.4 Menetelmän vahvuudet ja heikkoudet

Käytettävyystesteihin verrattuna heuristinen arviointi on halvempi ja nopeampi keino järjestelmän käytettävyyden arviointiin, sillä testit voidaan suorittaa ilman todellisia loppukäyttäjiä. Testauksessa ei myöskään tarvita erityisiä testilaitteistoja tai tiloja, kuten esimerkiksi käytettävyyslaboratoriossa tehtävissä käyttäjätesteissä. Menetelmän etuja ovat lisäksi, että se ei vaadi monimutkaista suunnittelua tai aikatauluttamista, ihmisiä on yleensä helppo motivoida tekemään arviointi ja menetelmää voidaan käyttää jo suunnittelun alkuvaiheessa [Nielsen & Molich 1990]. Testit on myös vaivatonta uusia projektin eri vaiheissa ja näin testejä voidaan käyttää esimerkiksi suunnittelun eri vaihtoehtojen priorisointiin.

Arviointi ilman todellista käyttäjää

Käyttäjän jättäminen testauksen ulkopuolelle on menetelmän ehdoton heikkous. Todellisten käyttäjien tarkastelu esimerkiksi käytettävyystesteissä paljastaa monesti yllätyksellisiä tilanteita ja ongelmakohtia, joita asiantuntija-arvioissa ei millään pystytä ennakoimaan. Useimmissa heuristisen arvioinnin säännöstöissä ei myöskään huomioi tehtäviä, joita järjestelmällä suoritetaan, jolloin käyttäjän kannalta ongelmallisia toimintoja voi jäädä huomioimatta. Jokaisen arvioijan ajatusmaailma heijastuu tuloksiin, jolloin tietyt virhetyypit voivat tulla ohitetuksi. Esimerkiksi järjestelmän suunnittelijat saattavat olla sokeita virheilleen.

Järjestelmän ja käyttäjän vuorovaikutus on monessa tapauksessa niin monitahoista, etteivät pelkistetyt menetelmät, kuten heuristinen arviointi, pysty huomioimaan tätä riittävästi [Cockton & Woolrych 2002]. Tuloksena on vääriä hälytyksiä käytettävyysongelmista. Esimerkiksi käytettävyysongelmaksi saatetaan raportoida tilarivin sekavasti muotoiltu ilmoitus, vaikka todellisuudessa käyttäjä ei välttämättä välitä lukea tilarivin tekstejä ollenkaan.

Heuristinen arviointi saattaa myös löytää enemmän vähäisiä ongelmia kuin merkittäviä ongelmia [Jeffries et al. 1991], joten ongelmien luokitteluun tulee kiinnittää huomiota. Heuristisella arvioinnilla saattaa myös jäädä jopa puolet ongelmakohtia huomaamatta [Law & Hvannberg 2002, Cockton & Woolrych 2002]. Menetelmällä ei myöskään välttämättä tarjoa minkäänlaista korjausehdotusta löydetuille ongelmille [Nielsen & Molich 1990]. Sen tarkoituksena on vain löytää ongelmakohtia, mutta se ei mittaa millään tavalla suunnittelun tavoitteiden saavuttamista [Kamper 2002]. Tämä voi johtaa virheellisiin korjausehdotuksiin.

Kustannussäästöt voivat olla riski

Suurin vaara tulee kustannussäästöjen hakemisesta. Heuristisessa arvioinnissa aikaa ja rahaa säästetään kaventamalla näköpiiriä, mistä käytettävyysongelmia etsitään. Kun arvioijat etsivät ongelmakohtia rankasti karsitun heuristiikkasäännöstön avulla arvioinnit sujuvat nopeasti, mutta ongelmakohtia voi jäädä huomaamatta. Heuristisen arvioinnin ongelmana on myös se, että arvioijat saattavat tutkia käyttöliittymän käyttämällä pelkästään maalaisjärkeä ongelmakohtien etsimisessä ja jättää heuristiset säännöt huomioimatta kokonaan. Cocktonin ja Woolrychin [Cockton & Woolrych 2002] tutkimuksien mukaan heuristisen arvioinnin tulokset riippuivat pitkälti arvioijan ammattitaidosta, ei niinkään heuristisista säännöistä. Heuristiikan avulla löydettiin vain vähäpätöisiä käytettävyysongelmia.

Heuristisen arvioinnin tuoma aikasäästö saattaa tietyissä tilanteissa olla harhaa. Erityisesti, jos arviointiin osallistuu useita arvioijia, on huomioitava myös tulosten

kokoamiseen ja analysointiin kuluva aika. Vaikka itse arvioinnit sujuisivatkin nopeasti vie muu käsittely paljon aikaa ja tällöin säästö esimerkiksi käytettävyystesteihin verrattuna kutistuu.

Tulosten vakuuttavuus

Vaikka käytetyllä käytettävyydenarviointimenetelmällä löydettäisiin paljon käytettävyysongelmia ei tuloksista ole hyötyä, jos päätöksiä tekevä taho ei vakuutu aineistosta. Heuristisen arvioinnin ongelmana on tässä näkökulmassa se, että arvioinnin suorittavat usein organisaation omat henkilöt ja tuloksena on kvalitatiivista aineistoa. Johtotasolla päätöksiä tehdään kuitenkin monesti mieluummin kvantitatiivisen aineiston perusteella [Brooks 1994]. Heuristisen arvioinnin tuloksia pidetään liian helposti vain muutaman henkilön mielipiteenä, kun taas käytettävyystesteistä saatu tilastoaineisto esimerkiksi suoritusajoista on helpompi hyväksyä faktana.

Käytettävyystestien etuna on myös se, että päättävien tahojen henkilöitä voidaan ottaa mukaan seuraamaan testejä tai jopa osallistumaan niihin. Kun he omin silmin näkevät käyttäjän painiskelemassa ongelmien parissa on jatkotoimenpiteiden perustelu helpompaa. [Karat 1994]

Kärkevintä kritiikkiä heuristista arviointia kohtaan ovat esittäneet Gilbert Cockton ja Alan Woolrych [Cockton & Woolrych 2002]. Heidän mielestään alennetuissa käytettävyyden arviointimenetelmissä (discount usability methods), joihin heuristinen arviointikin kuuluu, käytettävyysongelmia selvitetään huolimattomasti ja menetelmissä mennään siitä missä aita on matalin.

2.2 Käytettävyystestit

Käytettävyystesti (usability test) on käytettävyystutkimuksen perusmenetelmä, jossa järjestelmän ongelmakohtia pyritään löytämään todellisten käyttäjien kanssa

tehtävillä testeillä [Nielsen 1993, 165]. Se tarjoaa luotettavan keinon selvittää miten järjestelmää käytetään ja osoittaa selkeästi käytettävyysongelmia aiheuttavat kohdat.

Testeihin kutsutaan yleensä kolmesta viiteen käyttäjää [Nielsen & Landauer 1993]. Lisäksi ennen varsinaisia testejä itse testijärjestelyitä ja –tehtäviä on syytä testata yhden tai kahden käyttäjän kanssa, jotka eivät osallistu lopulliseen testiin. Nämä esitestit paljastavat usein puutteita ja ongelmia testitehtävissä, jolloin viat on mahdollista korjata ennen varsinaisia testejä [Nielsen 1993, 174].

Varsinaisiin testeihin osallistuvien käyttäjien lukumäärästä on esitetty myös vastakkaisia argumentteja. Nielsen ja Landauer perustelivat viiden käyttäjän riittävyden matemaattisilla laskelmilla, jotka johdettiin heidän toteuttamista käytettävyystesteistä saamistaan tilastoaineistosta. Näiden mukaan viidellä käyttäjällä pystyttiin löytämään jopa 85% järjestelmän käytettävyysongelmista. Wollrych ja Cockton taas huomasivat, että tulokset riippuivat testikäyttäjistä, testissä suoritettavista tehtävistä ja käytetyistä laitteista eikä viidellä käyttäjällä välttämättä saada selville löydettyjen ongelmien vakavuutta tai esiintymismäärää [Wollrych & Cockton 2001].

Testattavien valinnassa tärkeintä on, että he edustavat tuotteen todellisia käyttäjiä. Johtotason henkilöt saattavat olla hyvinkin innokkaita osallistumaan testeihin, mutta jos heidän työtehtäviinsä ei kuulu tuotteen käyttö, ei heitä pidä ottaa myöskään testihenkilöiksi. Valinnassa pitäisi myös huolehtia siitä, että testihenkilöt edustavat riittävän laaja-alaisesti käyttäjäkuntaa. Jokainen tuote pitäisi testata ainakin noviisikäyttäjillä ja mahdollisesti myös kokeneilla käyttäjillä [Nielsen 1993, 177].

2.2.1 Testien rakenne

Testissä käyttäjän tulee suorittaa ennaltamäärättyjä tehtäviä ja suoritusta tarkkailevat arvioijat, jotka pyrkivät löytämään käytettävyysongelmia käyttäjää tarkkailemalla. Käyttäjää pyydetään yleensä ajattelemaan ääneen testin aikana, jotta testaajat saavat selvemmän käsityksen mihin käyttäjän valinnat perustuvat. Ääneen ajattelu helpottaa tilanteen taltiointia ja myöhempää analysointia.

Testin toteutus voi vaihdella. Testejä voidaan tehdä käyttäjän luona, jolloin pyritään mahdollisimman autenttiseen tilanteeseen, tai erillisessä käytettävyyslaboratoriossa, jolloin testitilanne pyritään vakioimaan eri käyttäjien kesken.

Testitehtävien tulisi pyrkiä simuloimaan työtehtäviä, joihin järjestelmää tullaan oikeassakin tilanteessa käyttämään. Tehtävien tulisi myös riittävästi kattaa käyttöliittymän tärkeimmät ominaisuudet. Niiden tulisi olla riittävän yksikertaisia, jotta käyttäjä pystyy suorittamaan ne annetun aikarajan sisällä, mutta ei kuitenkaan täysin triviaaleja. Ensimmäisen tehtävän on kuitenkin hyvä olla varsin helppo, jotta käyttäjän alkujännitystä saadaan lievennettyä ja itsetuntoa kohotettua. Tavoitteena tulisi jokaisessa tehtävässä olla selkeän päämäärän saavuttaminen eikä vain yleinen järjestelmän selailu. Erityisesti viimeiseksi tehtävän kohdalla käyttäjän on tärkeä saada tuntee saavuttaneensa jotain. [Nielsen 1993, 185]

Testit rakentuvat esittelystä, varsinaisista testeistä ja palauteosiosta. Esittelyssä käyttäjä totutellaan testitilanteeseen ja harjoitutetaan ajattelemaan ääneen. Harjoittelua tarvitaan sillä harva käyttäjä pystyy muuten antamaan riittävän tarkkaa selvitystä toiminnastaan. Käyttäjille voidaan esittelyvaiheessa antaa aikaa omatoimisesti tutustua järjestelmään ja lukea järjestelmän mukana toimitettavaa dokumentaatiota [Karat et al. 1992]. Varsinaiset testit aloitetaan lukemalla tehtävä ääneen ja antamalla tehtävä myös kirjallisesti käyttäjälle. Tehtävä annetaan kirjallisena, jotta käyttäjän ei tarvitse muistaa tehtävää ulkoa vaan hän voi tarkistaa tavoitteen milloin tahansa testin aikana. Tämän jälkeen käyttäjä aloittaa tehtävän suorittamisen selostamalla ääneen ajatuksiaan. Tehtävän aikana testaajat eivät anna vinkkejä tai kommentoi suoritusta. Jos testiin osallistuu useampia testaajia, yksi toimii päätestaajana joka johtaa testin suoritusta ja kommunikoi käyttäjän kanssa. Muut testaajat eivät missään tilanteessa kommentoi testin kulkua, jotta käyttäjää ei häiritä.

Seuraavaan tehtävään siirrytään, kun edellinen on suoritettu tai käyttäjä ei muutaman minuutin mietinnän jälkeen pysty etenemään tehtävässä. Kun kaikki tehtävät on suoritettu tai aika loppuu, käydään vielä palauteosio. Käyttäjä voi palauteosiossa antaa palautetta testistä ja järjestelmästä sekä samalla rentoutua testin jännityksestä. [Hertzum & Jacobsen 2001]. Osioon voi kuulua palautekyselyn täyttäminen, mutta tämä tulee hoitaa ennen vapaata palautekeskustelua. Suullisesti käyttäjältä on hyvä kysyä kommentteja ja parannusehdotuksia järjestelmän kehitykseen.

Testaajat havainnoivat käytettävyysongelmia jo testauksen aikana, mutta jos testitilanne on nauhoitettu, analysointi voidaan suorittaa myös testien jälkeen. Havainnot kootaan testiraportiksi, jossa käytettävyysongelmat on lueteltu ja niiden vakavuutta on arvioitu.

2.2.2 Löydettyjen käytettävyysongelmien raportointi

Käytettävyystestausta suorittavat usein eri henkilöt kuin ketkä vastaavat löydettyjen ongelmakohtien korjauksesta. Testaajien ja suunnittelijoiden näkökulmissa on merkittäviä eroja [Jeffries 1994], joten on tärkeää, että testaajat pystyvät tarjoamaan riittävästi tietoa testeissä ilmenneistä ongelmakohdista. Kunnollinen testiraportointi on merkittävässä osassa, jotta suunnittelijat pystyvät ymmärtämään ja toteuttamaan ehdotetut muutokset oikealla tavalla.

Robin Jeffries [Jeffries 1994] kuvaa kahdeksan kohtaa, jotka arvioijan on syytä ottaa huomioon käytettävyysongelmien raportoinnissa.

1. Kuvaa ongelma ja ratkaisu erillään
2. Tarjoa perustelut ongelmalle ja ratkaisulle
3. Liitä arvio ongelman vakavuudesta
4. Pohdi avoimesti vaihtoehtoja
5. Arvioi huolellisesti kaikki ratkaisut, jotka vaativat uuden toiminnallisuuden lisäyksen sovellukseen
6. Ole tietoinen omista painotuksistasi, joita arvioijana tuot mukaan arviointiin
7. Yritä katsoa jokaista ongelmaa usealta eri kantilta

8. Arvioi ongelmaraporttia kokonaisuutena, karsien paikallisia optimointeja ja kompromisseja yleisemmillä ratkaisuilla jotka sopivat sovellukseen kokonaisuudessaan

Oheisten ehdotuksien läpikäynti jokaisen löydetyn käytettävyysongelman kohdalla vaatii arvioijalta enemmän aikaa, kuin täysin vapaa raportointi. Tavoitteena on kuitenkin parantaa raportoinnin laatua, jolloin virheellisiä korjausehdotuksia tulee karsittua ja korjausten toteuttajat saavat kuvan ongelman laadusta.

2.2.3 Käytettävyysongelmien luokittelu

Käytettävyystestissä löydettyihin ongelmiin voidaan käyttää osittain samoja menetelmiä, joita esiteltiin jo heuristisen arvioinnin yhteydessä luvussa 2.1.3. Samaan ideaan, käytettävyysongelmien jakamiseen kriittisiin ja vähemmän kriittisiin, perustuu myös erityisesti käytettävyystestejä varten suunniteltu ongelmien jaottelu yhdeksän kriteerin perusteella [Jacobsen et al. 1998]. Menetelmässä arvioijat käyvät itsenäisesti läpi testien videotallennuksia ja ongelmakohdat poimitaan, jos ne täyttävät jonkin seuraavista ehdoista:

1. käyttäjä ilmaisee tavoitteen, mutta ei pysty ratkaisemaan tehtävää kolmessa minuutissa
2. käyttäjä luovuttaa
3. käyttäjä ilmaisee tavoitteen, mutta joutuu käyttämään vähintään kolme toimintoa saavuttaakseen sen
4. käyttäjän ratkaisu on muu, kuin tehtävässä vaadittiin
5. käyttäjä ilmaisee yllättyneensä
6. käyttäjä ilmaisee negatiivisen mielipiteen tai sanoo jonkin olevan ongelma
7. käyttäjä tekee muutosehdotuksen
8. järjestelmä kaatuu
9. arvioija muodostaa ryhmästä aiemmin huomattuja ongelmia yhden uuden ongelman

Kriittiseksi käytettävyysongelmaksi määritellään ongelma, jonka vähintään yksi arvioija on huomannut ja joka rikkoo jotakin kohdista 1,2 tai 8.

Tämänkin luokittelun osalta voidaan kuitenkin kyseenalaistaa ovatko kaikki menetelmän avulla esiintulleet käytettävyysongelmat ongelmia myös todellisessa työtilanteessa. Menetelmä on erittäin hyvä apu testin vetäjälle ja se tuo selkeät säännöt ongelmakohtien tunnistamiseen. Samalla se saattaa kuitenkin alentaa käytettävyysongelman määrittelykynnystä, jolloin myös virheellisten hälytysten riski kasvaa.

2.2.4 Menetelmän vahvuuksia ja heikkouksia

Käytettävyystestien ehdottomana vahvuutena heuristiseen arviointiin verrattuna on käyttäjän mukanaolo. Menetelmän avulla on mahdollista löytää vakavia käytettävyysongelmia, jotka todella haittaavat käyttäjiä. Käytettävyystesteissä pystytään mittaamaan nopeus- ja tarkkuusarvoja, joita on hyvin vaikea heuristisen arvioinnin avulla saada luotettavasti kerättyä [Brooks 1994]. Ehdottomana vahvuutena on myös vaihtoehtoisten näkökantojen ja toimintamallien havainnointi, joita kehittäjäryhmän voi muutoin olla vaikea kuvitella. Ääneenajattelu on tässä keskeisessä asemassa, mutta käyttäjän sanomiset eivät saa liikaa ohjata käytettävyysongelmien syiden haussa. On tärkeämpää tarkkailla mitä käyttäjä oli tekemässä, kun käytettävyysongelma tapahtui, kuin luottaa käyttäjän arvioihin miksi ongelma syntyi [Nielsen 1993, 195]. Ääneenajattelu saattaa vaikuttaa myös käyttäjän suorituksiin testin aikana. Monille käyttäjille toiminnan samanaikainen selostaminen hidastaa itse tehtävän suoritusta, toisille taas ääneenajattelu helpottaa ongelmakohtien ratkaisua. Ennen varsinaisia testitehtäviä on siis hyvä harjoituttaa käyttäjää muutamalla ääneenajattelutehtävällä.

Testaajien ammattitaito vaikuttaa tulosten kelpoisuuteen eli siihen mittaako testi sitä mitä sen on haluttu mittaavan [Nielsen 1993, 179]. Tulosten analysointi on myös hyvin riippuvainen analysoijan taidoista ja käytettävyysperiaatteiden tuntemuksesta [Jacobsen et al. 1998]. Yksittäiseltä analysoijalta voi merkittävä osa käytettävyysongelmista jäädä huomioimatta, joten analysointi on syytä suorittaa

ryhmässä. Käytettävyysasiantuntijat, joilla on kokemusta tutkittavasta järjestelmästä, soveltuvat parhaiten tähän tehtävään.

Testattavien määrän karsiminen saattaa heikentää tuloksia. Eri käyttäjien henkilökohtaisissa ominaisuuksissa on huomattavia eroja, jotka heijastuvat myös testituloksiin. Gray ja Salzman [Gray & Salzman 1998] puhuvat villin kortin vaikutuksesta (wildcard effect), jossa pienessä testikäyttäjäryhmässä yksi taidoiltaan poikkeuksellinen käyttäjä heilauttaa tulosten keskiarvoa merkittävästi parempaan tai huonompaan suuntaan. Pienessä testikäyttäjäryhmässä voi tuloksiin myös jäädä merkittäviä puutteita. Jos joukkoon ei ole tullut yhtäkään esimerkiksi noviisikäyttäjää saattaa osa ongelmista jäädä kokonaan löytymättä ja löydettyjen ongelmien vakavuudesta saatetaan tehdä virheellisiä arvioita [Cockton & Woolrych 2002].

Kustannuslaskelmia tehdessä on myös muistettava, että testikäyttäjien määrä on vain yksi osa kulueristä. Suunnittelu ja analysointi vievät usein enemmän aikaa kuin itse testaus, joten testattavien määrän karsimisella ei välttämättä ole suurtakaan merkitystä kokonaiskustannuksissa.

Testattavien valinta vaikuttaa myös tulosten pätevyyteen [Nielsen 1993, 169]. Jos testikäyttäjät eivät edusta riittävässä määrin tuotteen oikeita käyttäjiä ei tuloksista voida tehdä päteviä johtopäätöksiä. Myös ennen varsinaisia testejä suoritettaviin esitesteihin valittavat testihenkilöistä ainakin toisen on tärkeää edustaa oikeita käyttäjiä.

Testitehtävien suunnittelussa ja valinnassa on oltava erittäin tarkkana. Helposti valittavaksi tulee vain tehtäviä, jotka paljastavat jo etukäteen ilmeisiä käytettävyysongelmia. Huolimattomasti valituissa tehtävissä saattaa käyttäjän kannalta merkittäviä työtehtäviä ja käyttötapoja jäädä tutkimatta, jolloin myöskään tulokset eivät heijasta todellista tilannetta. Varsinkin monimutkaisten käyttötilanteiden testaus on vaikeaa. Yleensä tehtävissä käyttäjät suorittavat vain yksittäisiä pieniä tehtäviä, kun taas todellisuudessa käyttäjä yhdistelee monia tehtäviä

jonkin suuremman tehtävän suorittamiseksi [Dicks 2002]. Tehtävien suunnittelua voi helpottaa esimerkiksi tarkkailemalla käyttäjien työskentelyä etukäteen ennen testien suunnittelua ja ottamalla näissä ilmenneitä työtehtäviä osaksi testiä.

Johdannossa esitellystä Gould:in ja Lewis:in käytettävyyden määritelmästä [Gould & Lewis 1983] on huomioitava kohdat ”hyödyllisyys” ja ”helppo käyttää”. Liian helposti käyttäjätesteissä keskitytään vain tutkimaan helppokäyttöisyyttä, mutta hyödyllisyys eli kysymykset ”tekeekö järjestelmä sen mitä sen pitikin tehdä” ja ”onko se yleisesti toimiva ja käytettävä” jäävät tutkimatta.

Itse testitilanteesta on huomioitava, että se ei koskaan vastaa täysin normaalia käyttötilannetta [Dicks 2002]. Käyttäjää häiritsee ja jännittää enemmän tai vähemmän tilanne, jossa hänen pitäisi suorittaa tehtäviä muiden henkilöiden tarkkaillessa ja mahdollisesti myös tallentaessa tilanteen. He saattavat yrittää miellyttää arvioijaa ja vähätellä ongelmien vakavuutta. Käyttäjälle onkin tärkeää painottaa, ettei testissä olla tutkimassa käyttäjän taitoja vaan mahdolliset ongelmat osoittavat puutteita vain testattavassa järjestelmässä. Jännityksen lievittämiseksi testit aloitetaan esittelyvaiheella, jolloin käyttäjät voivat rauhassa totutella tilanteeseen.

2.3 Käytettävyydenarviointimenetelmien vertailun teoria

Yksittäisen arviointimenetelmän tehokkuuden tutkiminen perustuu usein käytettävyyden määritelmästä johdettuihin mittareihin. Nielsen määrittelee käytettävyyden viiden osan summana [Nielsen 1993]:

- **Opittavuus:** järjestelmän tulisi olla helposti opittava, jotta käyttäjät voivat nopeasti keskittyä töiden tekemiseen eikä järjestelmän opiskeluun
- **Tehokkuus:** Järjestelmän tulisi olla tehokas eli kun käyttäjä on kerran oppinut käyttämään, hän voi saavuttaa korkean tuottavuuden
- **Muistettavuus:** Järjestelmän tulee olla helposti muistettava eli vaikka käyttäjä pitäisi järjestelmän käytöstä taukoa hänen on helppo tauon jälkeen palata käyttäjäksi ilman, että joutuu opettelemaan kaiken uudestaan.

- Virheet: Järjestelmässä tulisi olla alhainen virhesuhde niin, että käyttäjät eivät tee useita virheitä käytön aikana ja vaikka tekisivätkin niin niistä on helppo toipua. Vakavia virheitä ei saa sattua ollenkaan
- Tyytyväisyys: Järjestelmän käytön tulisi olla miellyttävää eli käyttäjä ovat subjektiivisen tyytyväisiä käyttäessään järjestelmää

Tehokkuus näkökulmaa voidaan tutkia esimerkiksi käytettävyytsteissä mittaamalla testitehtävän suoritukseen kuluva aikaa. Tyytyväisyyttä taas voidaan kysyä suoraan käyttäjältä testin jälkeisessä haastattelussa.

Heuristisen arvioinnin ja käytettävyytstestien tulosten vertailu ei kuitenkaan ole mutkaton tehtävä menetelmien poikkeavuuksien takia. Nielsenin määritelmästä johdetut mittarit soveltuvat lähinnä havainnointimenetelmien, kuten käytettävyytstestien, mittaukseen, sillä niissä tietolähteenä käytetään todellista käyttäjää. Luotettavia mittareita, joita voitaisiin käyttää kumpaankin menetelmään, on haasteellista laatia.

2.3.1 Löydettyjen käytettävyyso Ongelmien määrän vertailu

Yksi selkeä mittari on löydettyjen käytettävyyso Ongelmien määrä. Law:n ja Hvannbergin tutkimuksessa [Law & Hvannberg 2002] mittareina olivat (1) kuinka monta yksittäistä käytettävyyso Ongelmaa kumpikin menetelmä löysi ja (2) kuinka monta yhteistä käytettävyyso Ongelmaa menetelmillä löydettiin. Näissä vertailuissa myös luokittelutiedot otettiin mukaan ja tutkittiin kuinka monta vakavaa ja vähäistä käytettävyyso Ongelmaa menetelmillä löydettiin. Hartson tutkimusryhmineen [Hartson et al. 2003] lisää tähän listaan myös eroavaisuuksien vertailun eli kuinka monta käytettävyyso Ongelmaa menetelmä löysi, joita toisella menetelmällä ei löydetty.

Pelkkien löydettyjen käytettävyyso Ongelmien määrän huomioiminen menetelmiä vertailtaessa voidaan kyseenalaistaa. Heuristisessa arvioinnissa löydetään usein ongelmia, joita todellinen käyttäjä ei miellä käytettävyyso Ongelmiksi ollenkaan eli ns. virheellisiä hälytyksiä [Law & Hvannberg 2002]. Löydettyjen ongelmien

oikeellisuutta voidaan arvioida peilaamalla ongelmaa käytettyyn heuristiikkaan ja pyytämällä arvio usealta arvioijalta, mutta tämäkään ei ratkaise ongelmaa täysin [Hartson et al. 2003]. Molemmissa menetelmissä löydettyjen käytettävyysongelmien listalle saattaa tulla myös kohtia, jotka tarkemmassa arvioinnissa paljastuvat yhden käytettävyysongelman kaksoiskappaleiksi tai sisältävän osittain päällekkäisiä kohtia muiden käytettävyysongelmien kanssa [Gray & Salzman 1998]. Siten heuristisen arvioinnin ja käytettävyystestien vertailu löydettyjen käytettävyysongelmien perusteella ei ole täysin luotettava keino. Kaksoiskappaleiden ja virheellisten hälytysten karsimiseksi käytettävyysongelmalistat on syytä analysoida huolellisesti ennen vertailun tekemistä. Kaksoiskappaleiden löytämistä hankaloittaa kuitenkin se, että käytettävyysongelmat raportoidaan usein vapaamuotoisesti ja eri arvioijat saattavat kirjata löytämänsä ongelmat hyvinkin erilaisin sanakääntein.

Raportoinnin ongelmia voidaan helpottaa yhtenäisellä raportointilomakkeella. Jeffries tutkimusryhmineen [Jeffries et al. 1991] käytti eri menetelmien vertailussa lomaketta, joka täytettiin jokaisen löydetyn käytettävyysongelman kohdalla erikseen. Lomakkeella kuvattiin ongelma ja kerrottiin kuinka se löydettiin. Valmiit kysymykset lomakkeella saattavat ohjaata arvioijan vastauksia, joten lomakkeen muotoilussa täytyy olla hyvin tarkkana. Lomakkeen avulla voidaan kuitenkin helpottaa eri menetelmien vertailua, kun eri menetelmillä löydettyistä käytettävyysongelmista saadaan yhteneväiset tiedot.

2.3.2 Arviointimenetelmien vertailun haasteita

Hartson tutkimusryhmineen [Hartson et al. 2003] esittää käytettävyysmenetelmien vertailuun mittareiksi mm. tehokkuutta, jossa huomioidaan kuinka monta todellista ongelmaa löydettiin verrattuna kaikkiin raportoituihin ongelmiin ja kuinka monta todellista ongelmaa järjestelmässä oikeasti on. Tätä on kuitenkin jokseenkin mahdoton käyttää normaalitestitilanteessa, sillä ei voida millään selvittää kuinka monta todellista virhettä järjestelmästä kokonaisuudessaan on. Tällaiset mittarit soveltuvatkin vain teoriapainotteisiin tutkimuksiin, jossa testattava järjestelmä

rakennetaan pelkästään menetelmien vertailua varten ja siihen tehdään tahallaan käytettävyysongelmia.

Käytettävyysongelmien luokittelu tuo myös vertailuongelmia. Jos sekä heuristisessa arvioinnissa että käytettävyystesteissä löydettyjen käytettävyysongelmien luokittelun hoitavat samat henkilöt, voi aiemmin tehty luokittelu vaikuttaa jälkimmäisen luokittelun tuloksiin. Yksi ratkaisu tähän pulmaan on käyttää ns. sokeita luokittelijoita, jotka kukin osallistuvat vain yhden menetelmän tulosten luokitteluun eivätkä he saa tietoa itse testitilanteesta tai osallistujista. Heidän tehtävänä on vain arvioida löydösten vakavuus. Myös tämän jälkeen tehtävän kaikkien löydösten yhteenvedon laatimisessa tulisi kiinnittää huomiota arvioinnin luotettavuuskysymyksiin. [Gray & Salzman 1998]

Arviointimenetelmiä vertailtaessa yksi haasteista on vertailukelpoisen arviointitilanteen järjestäminen. Heuristista arviointia suorittaville asiantuntijoille ei tule antaa merkittävästi enempää aikaa järjestelmän arviointiin kuin mitä käytettävyystesteihin osallistuville käyttäjillekään. Arvioinnit ja testit tulisi suorittaa mahdollisimman samanlaisessa ympäristössä, mahdollisesti jopa samassa tilassa samoilla laitteilla. [Gray & Salzman 1998]

2.3.3 Kustannustehokkuus

Tulosten vertailun lisäksi arviointikeinoja voidaan järjestää kustannustehokkuuden mukaan. Mittareina voidaan tällöin käyttää testien kustannuksia ja käytettyä aikaa. Menetelmistä riippumatta suurin kustannuserä on usein käytetty aika. Muiden kustannusten arvioiminen ja tulosten luotettava vertailu onkin hankalampaa, sillä testilaitteistona saatetaan käyttää esimerkiksi ylimääräiseksi jäänyttä tietokonetta eikä testitilastakaan usein makseta mitään, jos testit ja arvioinnit tehdään esimerkiksi yrityksen omassa neuvotteluhuoneessa. Voikin olla täysin perusteltua käyttää kustannustehokkuuden mittarina vain käytettyä aikaa. Tällöin tulee kuitenkin ottaa

huomioon myös testien valmisteluun ja analysointiin sekä arvioijien ja testaajien kouluttamiseen kuluva aika testien suorittamisen vaatiman ajan lisäksi.

Jeffries tutkimusryhmineen [Jeffries et al. 1991] arvioi eri menetelmien hyötyjä ja kustannuksia vertailemalla menetelmällä löydettyjen käytettävyysongelmien vakavuutta käytettyyn aikaan. He pisteyttivät jokaisen ongelman sen vakavuuden mukaan asteikolla yhdeksästä yhteen (vakava sai 9 pistettä, vähäinen 1 pisteen) ja summasivat nämä pisteet. Näin jokaiselle menetelmälle saatiin vakavuussuhdeluku, jota verrattiin menetelmän läpikäyntiin kuluneeseen aikaan. Kulunutta aikaa laskettaessa huomioitiin sekä analyysiin että arviointimenetelmän ja tutkittavan järjestelmän opetteluun käytetty aika.

3 Sisällönhallintajärjestelmät

Internetin alkuaikoina WWW-sivut koostuivat yksittäisistä html-ohjelmointikielellä toteutetuista sivuista, jotka oli linkitetty toisiinsa. Näitä sivuja ylläpidettiin manuaalisesti tekemällä muutokset jokaiseen sivuun erikseen. Jokaista muutosta varten sivu oli ensin haettava esimerkiksi ftp-yhteyden avulla www-palvelimelta, muutokset ohjelmoitava html-kieltä käyttäen ja lopulta sivu palautettava takaisin palvelimelle. Tällaisia www-sivuja, joissa sisältö säilyy muuttumattomana ellei dokumenttiin tehdä manuaalisesti muutoksia, sanotaan staattisiksi www-sivuiksi.

Staattisten www-sivujen ylläpito on kuitenkin äärimmäisen työlästä, jos sivuja paljon tai päivityksiä tulee usein. Niinpä apuun ovatkin tulleet dynaamiset www-sivut, joiden sisältö rakennetaan jokaisella käyttäjän vierailukerralla uudestaan automaattisesti www-palvelimella. Esimerkkinä dynamisesta www-sivusta ovat keskustelufoorumit, joissa käyttäjien lähettämät viestit tallentuvat tietokantaan, josta www-palvelin osaa päivittää sivun sisällön näyttäen myös viimeisimmät viestit.

Sisällönhallintajärjestelmät ovat järjestelmiä, joilla voidaan hallita dynaamisia sivustoja. Sisällönhallintajärjestelmiä käytetään myös paljon muuhunkin kuin pelkästään www-sivujen ylläpitoon. Käyttökohteita ovat myös yrityksen sisäiset Intranetit ja vain rajatulle yrityksen ulkopuoliselle ryhmälle, esimerkiksi asiakkaille ja yhteistyökumppaneille, avoimet Extranetit. Monesti raja sisällönhallintajärjestelmien ja muiden järjestelmien välillä on häilyvä, sillä samaa järjestelmää voidaan käyttää esimerkiksi dokumenttienhallintaan, asiakashallintaan ja viestintäjärjestelmänä. Monet sisällönhallintajärjestelmät voidaan myös integroida hyvin tiiviisti muiden yrityksessä käytössä olevien liiketoimintajärjestelmien kanssa, joten rajanvedot eri käyttötarkoitusten välillä tulevat varsin keinotekoisiksi.

3.1 Rakenne ja käyttö

Sisällönhallintajärjestelmät on suunniteltu suurten tietomäärien hallintaan. Täten niiden tärkeimpiä ominaisuuksia ovat kunnollinen sisällön ja navigaation organisointi [Schrantz 1998]. Navigaatio ja sisältö onkin järjestelmissä erotettu toisistaan, mikä mahdollistaa vaivattomat muutokset rakenteeseen. Linkitys on mahdollista generoidaan automaattisesti, jolloin virheelliset viittaukset esimerkiksi poistettuihin sivuihin karsiutuvat.

Rakenteen lisäksi myös ulkoasu on mahdollista erottaa sisällöstä. Yleisenä ratkaisuna on koota erilaisia sivupohjia, joissa määritellään sivun ulkoasu. Nämä sapluunat eivät ota kantaa itse sisältöön vaan niissä on vain määritelty valmiita paikkoja, joihin sisältö haetaan tietokannasta. Näin ulkoasu voidaan vaihtaa sisällöstä riippumatta ja samaa sisältöä voidaan käyttää useissa eri yhteyksissä. Suuntauksena on myös yhä vahvemmin hyödyntää XML-tekniikoita¹, joilloin sisältö ja ulkoasu ovat täysin riippumattomia toisistaan.

Yhdellä web-julkaisujärjestelmällä voidaan hallita samanaikaisesti useita eri WWW-palveluita, esimerkiksi useita eri WWW-sivustoja. Käytettävyyden osalta tämä tuo haasteita, sillä ylläpitotoimintojen käyttöliittymät saattavat hyödyntää jokaisen palvelun omaa sivupohjaa. Tällöin ylläpitokäyttöliittymä vaihtelee eri www-sivustojen välillä liikuttaessa, vaikka koko ajan pysyttäisiinkin saman sisällönhallintajärjestelmän sisällä. Tätä ongelmaa ei kuitenkaan ilmene kaikissa web-julkaisujärjestelmissä, sillä osassa järjestelmistä ylläpitotoiminnoille on erillinen varsinaisesta WWW-palvelusta riippumaton käyttöliittymä.

Käyttöympäristöltään sisällönhallintajärjestelmät ovat ns. asiakas-palvelin – järjestelmiä (client-server system). Näissä yksi palvelin hoitaa usean käyttäjän

¹ Extensible Markup Language (<http://w3c.org/XML/>)

palvelun samanaikaisen. Rakenne on yleisesti käytössä verkkopalveluissa ja esimerkiksi www-sivut perustuvat tähän arkkitehtuuriin.

Käyttäjät jakautuvat erilaisiin käyttäjäryhmiin, joilla on eri tasoisia käyttöoikeuksia järjestelmään. Käyttöoikeuksien avulla voidaan rajoittaa ryhmien tai yksittäisten käyttäjien pääsyä ja muokkausmahdollisuuksia tietyille sivuille tai osioille. Käyttäjäryhmät vaihtelevat tavallisesta WWW-käyttäjästä, jolla on lukuoikeus vain rajattuun osaan järjestelmän sisältöön, aina järjestelmän pääkäyttäjään, jolla on kaikki mahdolliset oikeudet muokata rakennetta ja sisältöä sekä käyttöoikeuksia. Käyttöoikeuksiin liittyy myös monesti mahdollisuus työnkulun hallintaan. Tällöin esimerkiksi sisällön muokkaajan tallentamat muutokset eivät tule näkyviin ennen kuin ne on tarkistanut ja hyväksynyt kyseisen osion sisällöstä vastuussa oleva henkilö.

Sisällönhallintajärjestelmien käyttö eroaa monissa tilanteissa tavallisesta WWW-surffailusta [Dennis 1998]. Varsinkin ylläpitäjille järjestelmä on ensisijaisesti työväline, jota saatetaan käyttää päivittäin. Näin ollen käyttöliittymässä täytyy huomioida tehokäyttäjien vaatimukset mm. oikopolkujen muodossa. Toisaalta kaikkien järjestelmän toimintojen ei välttämättä tarvitse olla noviisikäyttäjälle suunnattuja vaan voidaan olettaa, että työkseen järjestelmää käyttävät ovat myös valmiita opettelemaan toimintojen käyttöä. Monimutkaisemmissa toiminnoissa on kuitenkin huomioitava, että käyttöopastus tai dokumentaatio on mahdollisimman helposti saatavissa ja eri toiminnot käyttävät yhdenmukaisia käyttöperiaatteita.

3.2 Sisällönhallintajärjestelmän etuja

Käyttäjien kannalta mieluista sisällönhallintajärjestelmissä on pyrkimys piilottaa tekniikka pinnan alle. Helpoimmillaan ylläpito muistuttaa yhdistelmää normaalista internet-selailusta ja tekstinkäsittelyohjelman käytöstä. Ylläpitäjällä ei tarvitse olla ohjelmointikokemusta tai tietämystä web-tekniikoista kuten HTML- tai Java-kielistä. Sisällön muokkaus tapahtuu internet-selaimen avulla ja joissakin sisällönhallintajärjestelmissä ylläpitotoiminnot on myös mahdollista hoitaa

ulkopuolisen ohjelmiston avulla esimerkiksi Word-tekstinkäsittelyohjelmalla tai WWW-sivujen ylläpitoon tarkoitettulla Macromedian Dreamweaver -ohjelmalla hyödyntäen WebDAV-protokollaa² [Whitehead 1998]. Näin käyttäjän ei tarvitse opetella uuden työkalun käyttöä.

Integrointimahdollisuudet eri järjestelmien kanssa helpottavat myös käyttäjän työtä. Parhaimmillaan dokumentin tallentaminen yhteen paikkaan tuo sen automaattisesti näkyviin niin intranetiin, yrityksen WWW-sivuille, extraneteihin ja muihin haluttuihin järjestelmiin. Muutokset tarvitsee tehdä vain yhteen paikkaan ja järjestelmä hoitaa loput. Järjestelmissä voi olla myös ominaisuuksia, joilla sisältö automaattisesti muokataan sopivaksi erilaisille laitteille. Näin samaa sisältöä voidaan katsella WWW-sivuilla, digi-tv:tä varten tekstistä voidaan näyttää esimerkiksi vain johdanto ja mobiililaitteita varten siitä on voitu karsia myös kuvamateriaali paremmin pienelle näytölle sopivaksi.

3.3 Sisällönhallintajärjestelmien käytettävyyden arviointi

Sisällönhallintajärjestelmät ovat yleisesti varsin monimutkaisia järjestelmiä. Käyttäjien tehtävät vaihtelevat suuresti sisällön julkaisusta aina rakenteen sekä ulkoasun määrittelyyn ja yhdellä käyttäjällä voi olla useita käyttörooleja. Järjestelmän käytettävyyden arvioinnin kannalta onkin oleellista tunnistaa eri roolit ja niiden tarpeet, jotta parannusehdotukset voidaan laatia tarpeiden mukaan.

Sisällönhallintajärjestelmiä hallitaan usein WWW-selaimen avulla ja käyttöliittymät yleisesti hyödyntävät WWW-sivuilta tuttuja käyttöliittymäelementtejä. Silti niitä voidaan pitää enemmän sovelluksina kuin perinteisinä WWW-sivuina. Sisällönhallinnan työkalut muun muassa muistuttavat hyvin läheisesti tekstinkäsittelyohjelman toimintoja. Toisaalta perinteisiin paikallisiin tietokonesovelluksiin verrattuna WWW-selaimet mahdollistavat toiminnoiltaan

² World Wide Web Distributed Authoring and Versioning

selvästi rajoitetumpia käyttöliittymiä. Kun tämän yhdistää järjestelmän monimutkaisuuteen on seurauksena helposti ongelmia käytettävyydessä. Varsinaiseen käytettävyyden arviointiin luokitteluongelmat sovelluksien ja WWW-sivustojen välillä eivät vaikuta, mutta korjausehdotusten tekoon se tuo ongelmia. Esimerkiksi painikkeiden ja linkkien käytössä suositukset sovelluksien ja WWW-sivustojen kanssa menevät ristiin. Paikallisissa sovelluksissa ei linkkejä käytetä juuri missään tilanteessa, mutta WWW-sivustoilla painikkeita tulisi käyttää vain, jos ollaan suorittamassa jokin toiminto. Linkkejä tulisi WWW-sivuilla hyödyntää siirtymiseen eri toimintojen välillä, kun taas sovelluksissa painikkeet käyvät tähänkin tarkoitukseen.

3.3.1 Järjestelmään lukittautumisen vaikutukset

Avoimissa järjestelmissä, kuten internet-kauppasivustoissa, on tärkeää pystyä poistamaan kaikki vakavat käytettävyysongelmat. Käyttäjän on äärimmäisen helppoa vaihtaa järjestelmää, jos hän törmää ongelmaan [Sutcliffe 2002]. Näissä järjestelmissä yhden käytettävyyden arviointimenetelmän käyttö ongelmien selvittämisessä ei ole riittävää, sillä riski joidenkin ongelmien ohittamisesta on liian suuri. [Cockton & Wollrych 2002]

Tässä työssä keskitytään järjestelmien ylläpitäjien näkökulmaan jolloin sisällönhallintajärjestelmissä tilanne on hieman erilainen. Järjestelmään investointi on erittäin suuri prosessi asiakkaalle ja kynnys vaihtamiseen on tämän jälkeen suuri, sillä sisällön siirtäminen toiseen järjestelmään on usein todella vaikeaa ellei mahdotonta ja ylläpitäjät on koulutettu tämän tietyn järjestelmän käyttöön. Asiakkaalle siis järjestelmän vaihtaminen tulee todella kalliiksi ja asiakasta voidaan pitää järjestelmään lukittautuneena [Shapiro & Varian 1999].

Sisällönhallintajärjestelmän toimittajan näkökulmasta katsottuna ei ole välttämätöntä karsia järjestelmästä kaikkia käytettävyysongelmia. On kuitenkin huomioitava, että käytettävyysongelmat vaikeuttavat myös toimittajan mahdollisuuksiin saada uusia

asiakkaita. Yritykset vaihtavat kokemuksiaan eri järjestelmien käytöstä ja huonot kokemukset mustaavat järjestelmätoimittajan maineen helposti. Tarjouskilpailuihin kuuluu usein järjestelmän käytön esittely. Erot kilpailevien järjestelmien välillä ovat pieniä, jolloin pienetkin ongelmakohdat saattavat ratkaista kilpailun. Tällöin aiemmat käytettävyydesteissä tehdyt kustannussäästöt saattavat kääntyä erittäin kalliiksi.

3.4 Nas-sisällönhallintajärjestelmä

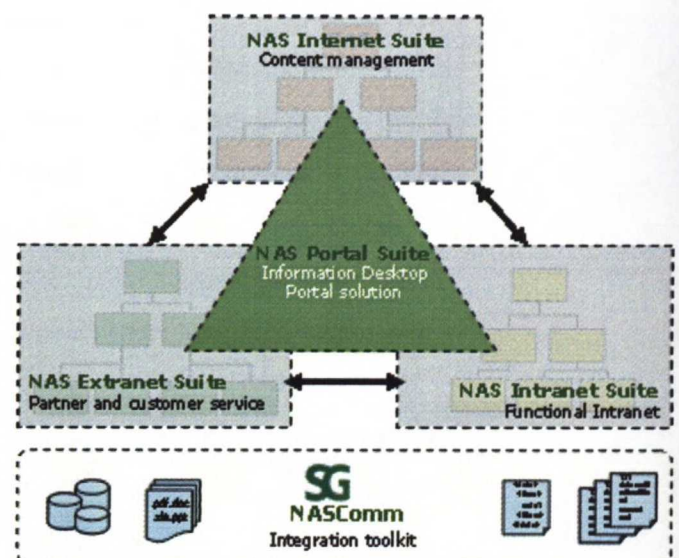
NAS-sisällönhallintajärjestelmän juuret vievät vuoteen 1998, jolloin Nedecon Oy julkaisi ensimmäisen selainpohjaisen version järjestelmästä. Yritysfuusioiden kautta ohjelmisto kulki Endero Oyj:lle, kunnes vuonna 2002 MBO-kaupalla ohjelmiston kehitys eriytettiin Systems Garden Oy:lle. Järjestelmällä on toteutettu noin 180 ratkaisua 120 eri asiakkaalle [Paronen 2004a] ja sillä on useita satoja tuhansia käyttäjiä, joista pääosa Suomessa ja Brasiliassa.

3.4.1 Järjestelmän rakenne

NAS-järjestelmän rakenne on modulaarinen eli järjestelmän toimintoja voidaan muokata asentamalla erillisiä moduuleja perusjärjestelmän ympärille. Esimerkkinä moduuleista ovat mm. keskustelufoorumi- ja uutispalsta-toiminnot.

Erilaisista

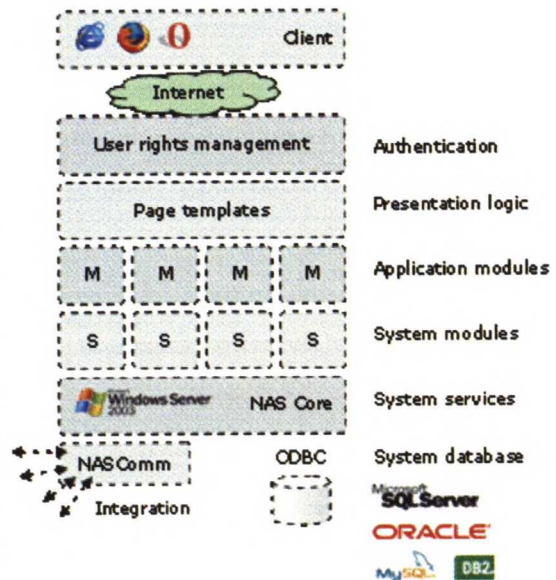
moduuliyhdistelmistä on tuotteistettu neljä tuotetta. NAS Internet Suite, NAS Extranet Suite ja NAS Intranet Suite sisältävät vain osan toiminnoista ja ovat suunnattu vain nimensä mukaiseen käyttötarkoitukseen. NAS Portal Suite on



Kuva 1. NAS-tuotteet

kokonaisratkaisu, jolla voidaan toteuttaa yrityksen sisäiset ja ulkoiset tiedonhallintaratkaisut. Se sisältää kolmen muun NAS-tuotteen toimintojen lisäksi myös ns. virtuaalityöpöytätoimintoja, joilla esimerkiksi sähköposti- ja kalenteriohjelmaa voidaan käyttää selaimen avulla. NAS-järjestelmän yhteydet muihin järjestelmiin on mahdollista toteuttaa NASComm-yhteyskäytävän avulla. Näin esimerkiksi tuotannonohjausjärjestelmät kuten SAP, voidaan yhdistää NAS-järjestelmään.

Järjestelmän arkkitehtuuri perustuu usealle päällekkäiselle tasolle. Pohjalla on Windows-ympäristössä toimiva ASP-kielellä toteutettu ydin, joka käsittelee käyttäjiltä tulevat kutsut ja ohjaa ne esimerkiksi tietokantaan ja palauttaa tulokset takaisin käyttäjälle. Ytimen päällä toimivat eri moduulit, jotka tarjoavat varsinaiset toiminnallisuudet. Järjestelmämoduulit (system modules) ovat järjestelmän perustoimintoja, esimerkkinä



Kuva 2. NAS-arkkitehtuuri

sisällönmuokkaustyökalu, ja näiden lisäksi asiakkaan tarpeen mukaan järjestelmää laajennetaan laajennusmoduuleilla (application modules) kuten keskustelufoorumi- tai gallup-moduuleilla. [Paronen 2004b]

Palvelun ulkoasu määräytyy sivupohjan avulla (page template), jossa määritellään paikat navigaatiolle ja sisällölle. Sivupohjat ovat yhdistelmä WWW-ohjelmointia (esimerkiksi HTML- tai XML-kieltä) sekä NAS-järjestelmän omaa ohjelmointikieltä, ns. metakoodia. Metakoodin avulla sivupohjaan voidaan esimerkiksi tehdä if-lauseita,

joilla ulkoasua voidaan vaihdella toiminto- tai sivukohtaisesti. Yhdessä sivustossa voi olla käytössä myös useita sivupohjia. Sivupohjassa määritellään erikseen paikat, mihin varsinainen sisältö kutsutaan. Sisältö voi olla esimerkiksi moduulin tuottama näkymä, kuten uutispalsta, tai tavallinen tekstisisältö.

Ylimmällä tasolla on käyttöoikeuksien hallinta. Jokainen käyttäjä todennetaan ennen järjestelmän toimintojen tarjoamista. Järjestelmässä käyttöoikeuksia voidaan asettaa osasto-, ryhmä- ja käyttäjäkohtaisesti. Jokainen käyttäjä kuuluu yhteen osastoon, jotka voivat esimerkiksi muodostua suoraan yrityksen organisaation mukaan. Lisäksi käyttäjä voi kuulua yhteen tai useampaan ryhmään. Ryhmät ovat kuin rooleja esimerkiksi pääkäyttäjä, joilla tietylle ryhmälle voidaan antaa tai poistaa oikeuksia osastorajoista riippumatta.

Käyttöoikeustasoja on myös useita lähtien kaikkien oikeuksien puuttumisesta ja täydelliseen rakenteen ja sisällön hallintaan. Jos käyttäjällä ei ole oikeutta sivuun tai toimintoon, järjestelmä piilottaa kyseisen elementin käyttäjältä kokonaan. Näin esimerkiksi sivustokartassa käyttäjä näkee linkit vain niille sivuille, joille hänellä on jonkinlaiset käyttöoikeudet.

3.4.2 Järjestelmän käyttö

NAS-järjestelmää käytetään tavallisella internet-selaimella. Selaimen vasemmassa reunassa on näkyvissä ylläpitotyökalut, joilla tärkeimpiä toimintoja ohjaillaan. Halutessaan käyttäjä voi myös piilottaa painikkeet.

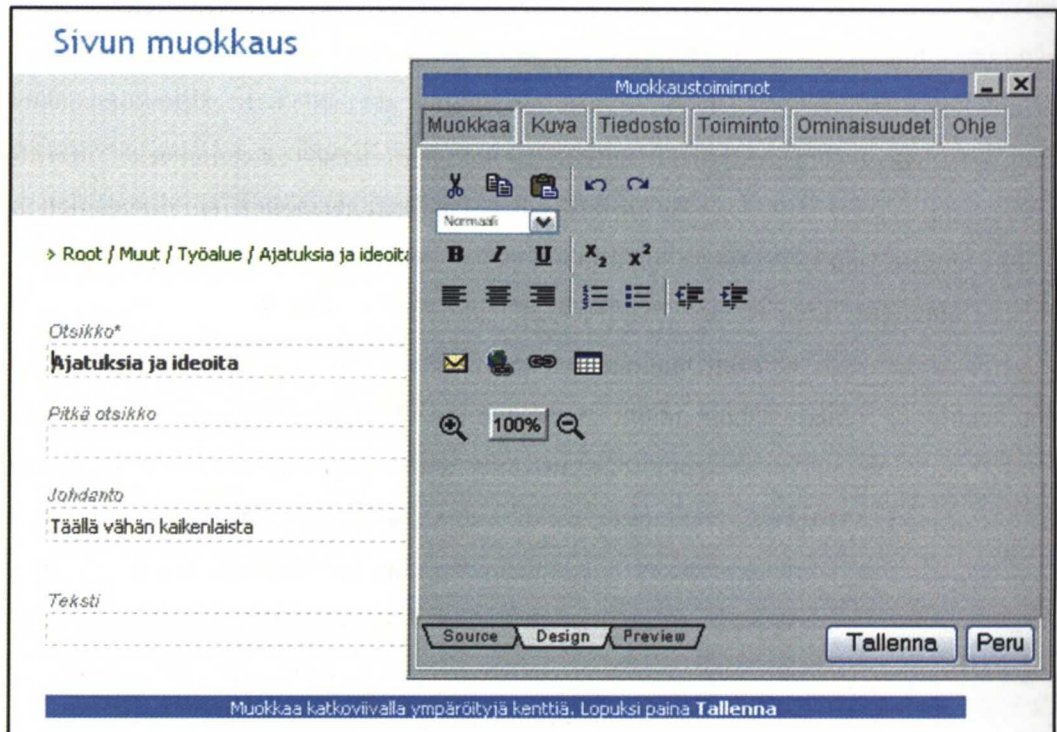
Valikon painikkeet on lajiteltu käyttöoikeustasojen mukaan. Pelkillä lukuoikeuksilla käyttäjä voi valikon



Kuva 3. Ylläpitotyökaluvalikon näyttömuodot

kautta muokata käyttäjätietojaan ja kirjautua järjestelmästä ulos. Tavallisessa www-järjestelmässä, joissa käyttäjien ei tarvitse kirjautua, nämä toiminnot on peruskäyttäjältä piilotettu kokonaan. Sivujen ja rakenteen muokkaus tapahtuu myös ylläpitotyökaluvalikon kautta. Käyttöoikeuksien ja toimintojen hallintaan liittyvät painikkeet ovat näkyvillä vain järjestelmän pääkäyttäjille.

Sivun muokkaus tapahtuu erillisen WYSIWYG-editorin (What You See Is What You Get) avulla. Editorin käyttö vastaa tavallista tekstinkäsittelyohjelmaa ja nimensä mukaisesti muutokset näkyvät heti ruudulla lopullisessa ulkoasussaan. Editorin avulla voidaan muokata tekstiä ja asettaa linkkejä, mutta myös kuvia ja liitetiedostoja. Sillä voidaan myös asettaa laajennusmoduulien tarjoamia toimintoja kuten uutispalstoja tai keskustelufoorumeita.



Kuva 4. Sivu editointitilassa WYSIWYG-editori avoimena

3.4.3 Järjestelmän ylläpitäjät

Valtaosa NAS-järjestelmien käyttäjistä on keskikokoisia tai suuria yrityksiä. Tällaisissa yrityksissä NAS:lla toteutettujen palveluiden ylläpidosta vastaa usein erillinen viestintä- tai tietotekniikkaosasto. Järjestelmän pääkäyttäjän oikeudet annetaan yrityksen koosta ja organisaatiosta riippuen yhdelle tai muutamalle henkilölle, jotka ovat yleensä osaston vetäjiä. He hoitavat käyttöoikeuksien hallinnan ja huolehtivat myös järjestelmän rakenteen pääpiirteistä.

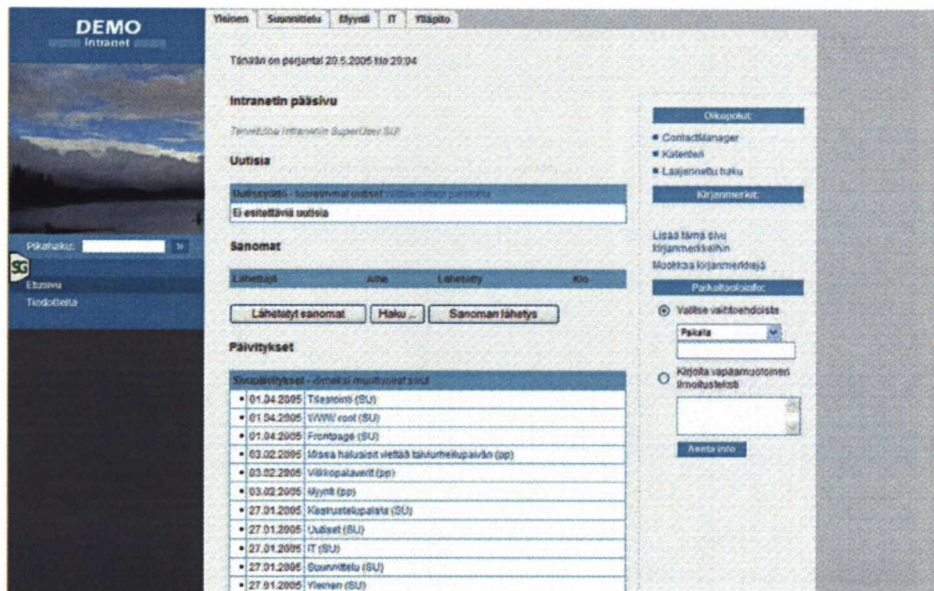
Järjestelmän pääkäyttäjä ei monesti vastaa sisällön hallinnasta ellei päävastuuta kannan viestintäosasto. Yksittäisiltä sivuilta pääkäyttäjä voi vastata lähinnä uusien toiminnallisuuksien asentamisesta, kuten uuden uutispalstan luomisesta. Myös rakenteeseen tehtävät suuremmat muutokset, kuten uuden osion luominen tai poistaminen voi kuulua pääkäyttäjän tehtäviin.

Itse sisällön vastuut voivat jakautua esimerkiksi organisaation mukaan, jolloin jokaiselta organisaation päähaaralta nimetään henkilö vastaamaan tiettyjen palvelussa olevien osioiden ylläpidosta. Näillä henkilöillä on oikeudet tehdä muutoksia sekä rakenteeseen, että sisältöön tässä vastuuosiossaan. Lisäksi varsinkin suuremmissa organisaatioissa voidaan nimetä vielä joukko ylläpitäjiä, joilla on oikeudet vain tehdä muutoksia tiettyihin olemassa oleviin sivuihin, mutta ei rakenneoikeuksia eli he eivät saa luoda tai poistaa sivuja. Pienemmissä yrityksissä koko palvelun hallinnasta voi vastata vain yksi henkilö.

4 Testien rakenne ja toteutus

4.1 Testisivusto

Heuristisia arviointeja ja käytettävyydestejä varten rakennettiin kuvitteellinen intranet-sivusto. Rakenteen pääpiirteiltään ja toiminnallisuuksiltaan sivusto vastasi useille Systems Gardenin asiakkaille toimitettuja intranet-sivustoja.



Kuva 5. Testisivustossa yhden testiosion pääsivu

Yksittäisen testiosion rakenne oli seuraava:

- Yleinen (Etusivu)
- Suunnittelu
 - Projektit
- Myynti
 - Uutiset
 - Keskustelupalsta
- It
 - Tiedotteita
- Ylläpito

Käyttäjän kirjautuessa järjestelmään, hän ohjautui automaattisesti yleinen-sivulle, joka toimi järjestelmän etusivuna.

Jokaista heuristista arvioijaa ja käytettävyydestä varten tehtiin identtiset kopiot rakenteesta ja sisällöstä. Näin jokainen testaaaja pääsi suorittamaan testitehtävät muista riippumatta eivätkä heidän tekemänsä muutokset vaikuttaneet muiden testaajien työhön. Jokaisella testaaajalla oli tehty omat käyttäjätunnuksensa, jotka ohjasivat testaajan kirjautumisen yhteydessä omalle osiolleen.

Osoimisen etuna oli myös se, että jokaisen testaaajan tekemät muutokset jäivät talteen ja niihin voitiin testien analysointivaiheessa tarvittaessa palata. Tätä osiota pystyttiin käyttämään muistiinpanojen tukena.

4.2 Testitehtävien suunnittelu

Tässä työssä arviointeja tuettiin valmiiksi määritetyillä tehtävillä, joita testaajien tuli suorittaa. Heuristinen arviointi voitaisiin suorittaa myös ilman tehtäviä, jolloin järjestelmä arvioitaisiin näkymä kerrallaan. Tässä työssä keskitytään kuitenkin vain tiettyjen käyttäjäryhmien käyttötarpeisiin, joten valmiilla testitehtävillä arviointi saatiin rajattua järkevästi. Samojen tehtävien käyttö sekä heuristiseen arviointiin että

käytettävyydesteihin mahdollisti myös vaivattomamman menetelmien tulosten vertailun.

Testitehtävien valinnasta Cordes [Cordes 2001] esittää kaksi vaihtoehtoa. Tuotteen tukemat tehtävät (product-supported task) valitaan nimensä mukaan tuotteen toimintojen mukaan. Näissä tehtävissä riskinä on, että tutkimatta jää, soveltuuko tuote käyttötarkoitukseensa eli vastaako tuote käyttäjän vaatimuksia. Toisena vaihtoehtona ovat käyttäjän määrittämät tehtävät (user-defined tasks), joissa käyttäjät itse määrittävät mitä tehtäviä he haluavat tuotteella suorittaa. Tämän lähestymistavan ongelmana on se, että tehtävien laatu voi vaihdella merkittävästi, kun osa testaajista haluaa suorittaa vain rutiinitehtäviä.

Tässä työssä testitehtävät valittiin NAS:n tukemien tehtävien mukaan, mutta näkökulmana valinnassa oli sisällön ylläpitäjän ja pääkäyttäjän työtehtävät. Se miksi käyttäjien ei annettu itse määritellä tehtäviä perustui siihen, että ensisijaisena tämän tutkimuksen tavoitteena on vertailla eri käytettävyyden arviointimenetelmiä. Jotta eri menetelmillä saatuja tuloksia voidaan luotettavasti vertailla, on tehtäväjoukon pysyttävä samana kaikissa testitilanteissa.

Testitehtävät on lueteltu liitteessä C. Käytettävyydesteissä käytettiin lisäksi erillisiä harjoitustehtäviä, joiden avulla käyttäjä pystyi harjoittelemaan ääneen ajattelua ja pahinta alkujännitystä voitiin näin myös helpottaa. Näiden testitehtävien kohdalla ei kerätty käytettävyyso ongelmia vaan ne toimivat vain harjoitteluosiona. Harjoittelutehtävät on lueteltu liitteessä B.

Tehtävien laadinnassa pyrittiin mukailemaan kuvitteellisen intranet-järjestelmän pääkäyttäjän mahdollisia työtehtäviä. Tarkempi vastaavuus todellisiin työtehtäviin olisi saatu käymällä ennen tehtävien laadintaa tarkkailemassa pääkäyttäjiä työssään tai suorittamalla haastattelukierros, mutta oma arvioni on ettei tätä voida pitää merkittävänä virhelähteenä. Tehtäviä laadinnassa oli mukana kaksi Systems Gardenin

työntekijää, joilla kummallakin on usean vuoden kokemus tämän alan projektien läpiviennistä ja asiakkaiden toimenkuvasta. He pystyivät varsin tarkasti kuvailemaan ylläpitäjien ja pääkäyttäjien työkuva. Myös käytettävyyystesteihin osallistuneet pitivät testitehtäviä todenmukaisina, kun heiltä asiasta kysyttiin palauteosion lopuksi.

Tehtävät laadittiin pääasiassa niin, että edellisen tehtävän epäonnistuminen ei vaikuta seuraavan tehtävän suoritukseen. Vain viides tehtävä vaati edellisessä tehtävässä luodun sivun muokkausta, mutta tässäkin muokkaustoimet pystyi tehdä jollekin muulle sivulle, jos edellisen tehtävän sivunluonti ei ollut onnistunut. Tällaisia tilanteita ei testeissä ilmennyt. Tehtävien valinnassa ja järjestyksessä huomioitiin myös se, että sekä alkuun että loppuun pyrittiin sijoittamaan hieman helpompia tehtäviä. Näin testaajan luottamusta itseensä pyrittiin vahvistamaan testin alussa ja lopuksi koitettiin saada tunne onnistumisesta.

4.3 Heuristisen arvioinnin suunnittelu ja toteutus

Heuristinen arviointi suoritettiin ennen käytettävyystestejä, jotta käytettävyystestien tulokset eivät vaikuttaisi arviointiin. Arviointiin osallistui allekirjoittaneen lisäksi kaksi henkilöä, mikä vastaa Nielsenin suosituksia [Nielsen 1992]. Kaikilla arvioijilla oli kokemusta järjestelmän käytöstä ja he osallistuvat myös järjestelmän kehitykseen. Kahdella arvioijista oli myös käytettävyyssasiantuntemusta ja kokemusta heuristisen arvioinnin toteutuksesta. Nämä kaksi ohjasivat myös myöhemmin pidettäviä käytettävyystestejä.

Ennen varsinaista arviointia kokoonnuimme johdantotilaisuuteen, jossa selvitin yleispiirteisesti arvioinnin tarkoituksen ja lyhyen johdatuksen heuristiseen arviointiin. Samalla jaettiin myös liitteessä A esitelty ohjeistus arviointiin sekä luvussa 2.1.1 esitelty Nielsenin ja Molichin kymmenen kohdan heuristiikka.

Johdannon jälkeen jokainen arvioija siirtyi omalle työpisteelleen suorittamaan arviointia. Ohjeistuksessa lueteltujen tehtävien avulla järjestelmää käytiin läpi ja

mahdolliset käytettävyysongelmat kirjattiin ylös. Jokaisen käytettävyysongelman kohdalla kirjattiin ylös ongelman kuvaus, mihin tehtävään se liittyi, mitä heuristiikan kohtaa tai kohtia se rikkoi ja ehdotus korjaukseksi. Jokainen arvioi myös alustavasti ongelman vakavuutta ohjeistuksessa kuvatulla asteikolla, joka on esitelty liitteessä A.

Kun jokainen arvioija oli saanut läpikäytyä kaikki tehtävät, keräännettiin yhteiseen palaveriin, jossa kaikki löydetyt käytettävyysongelmat käytiin läpi. Jokaisen löydöksen kohdalla arvioitiin uudelleen käytettävyysongelman vakavuutta ja päällekkäisiä löydöksiä yhdisteltiin.

4.4 Käytettävyystestien suunnittelu ja toteutus

4.4.1 Testikäyttäjien valinta

Käytettävyystesteihin osallistujat valittiin Systems Gardenin nykyisten asiakkaiden joukosta sekä yritysten, joille NAS-järjestelmää oltiin tarjoamassa joulukuussa 2004. Kutsut lähetettiin sähköpostitse sellaisille henkilöille, jotka tiedettiin toimivan NAS-järjestelmän pääkäyttäjinä tai ylläpitäjinä. Potentiaalisten asiakkaiden kohdalla kutsut lähetettiin niille, joiden arveltiin toimivan pääkäyttäjinä tai ylläpitäjinä, jos NAS otettaisiin kyseisessä yrityksessä käyttöön. Näin pyrittiin varmistamaan, että testikäyttäjät vastaavat järjestelmän todellisia käyttäjiä, mikä Drayn ja Siegelin mukaan on erittäin tärkeää testitulosten luotettavuuden kannalta [Dray & Siegel 1999].

Testikäyttäjien määräksi valittiin kymmenen käyttäjää. Nielsenin ja Landauerin laskelmissa [Nielsen & Landauer 1993] testaaminen yli viidellä käyttäjällä ei ole enää parantanut tuloksia oleellisesti, mutta halusin mukaan myös käyttäjiä, joilla ei ollut aiempaa kokemusta NAS:in käytöstä. Nielsenkin oli maininnut tärkeäksi sen, että testiin osallistuu sekä noviiseja, että kokeneita käyttäjiä [Nielsen 1993]. Yli viiden testikäyttäjän puolesta puhuttiin myös useissa papereissa [Wollrych & Cockton 2001, Dray & Siegel 1999].

4.4.2 Testin rakenne

Käytettävyystestit päätettiin jakaa neljään osaan suosituksen mukaisesti: alkuhaastatteluun, harjoitustehtäviin, varsinaisiin testitehtäviin ja loppuhaastatteluun [Karat 1994, Nielsen 1994b]. Alkuhaastattelun tarkoituksena oli ensisijaisesti lievittää testikäyttäjän jännitystä muutamalla vapaamuotoisella kysymyksellä, mutta näiden avulla pyrittiin myös profiloimaan käyttäjän kokemusta sisällönhallintajärjestelmistä ja yleisesti internetin käytöstä.

Tekstinkäsittelyohjelmien käyttökokemuksesta päätettiin kysyä, sillä NAS-järjestelmässä mm. wysiwyg-editorissa hyödynnetään hyvin paljon Office-ohjelmien mukaisia ikoneja.

Alkukysymyksillä haluttiin myös selvittää, kuuluuko käyttäjä todellakin rajattuun käyttäjäryhmään eli toimiiko tai toimisiko hän sisällön ylläpitäjänä tai pääkäyttäjänä. Tämän takia kysyttiin tehtävänimikettä sekä kokemusta sisällöntuotantojärjestelmistä ja ylläpitotehtävistä.

Alkukysymyksien jälkeen harjoiteltiin itse testitilannetta kolmella harjoitustehtävällä. Näiden ensisijainen tarkoitus oli harjoituttaa käyttäjä ajattelemaan ääneen. Testitehtävien avulla käyttäjä myös tutustui testisivustoon. Tehtävien lomassa hänelle myös selvitettiin testisivuston sivusto- ja navigaatorakenne, sillä nämä asiat eivät kuulu NAS:in toiminnallisuuksiin ja haluttiin rajata testin ulkopuolelle.

4.4.3 Testitilanne

Käytettävyystesteissä minua avusti Systems Gardenin työntekijä Sanna Karkkunen. Työnjako oli tällöin selkeä: minä hoidin kommunikoinnin testikäyttäjän kanssa ja Sanna kirjasi testin tapahtumat ylös. Muistiinpanoihin pyrittiin kirjaamaan yleisesti käyttäjän toimet ja sanomiset sekä onnistuiko tehtävän suoritus. Testihenkilöiden reaktioita, kuten hermostumisen ja hämmennyksen osoittamista, ja mahdolliset apupyynnöt pyrittiin kirjaamaan ylös. Koko testitilanne myös videoitiin. Testit suoritettiin pääasiassa erillisessä huoneessa käyttäjän omalla työkoneella tai

neuvotteluhuoneessa. Käyttäjille lähetettiin NAS:in käyttöohje etukäteen ja ohjekirja oli saatavilla myös testi aikana.

4.4.4 Pilottitesti

Ennen varsinaisia käytettävyystestejä suoritimme Systems Gardenin toimistolla pilottitestin, jolla testasimme testitehtävien ja -kysymyksien toimivuutta sekä mittasimme testin vaatimaa aikaa. Pilottitestiin testikäyttäjäksi saimme Systems Gardenin uuden työntekijän. Hän oli ensimmäistä päivää töissä, joten kokemusta NAS:in käytöstä ei ollut vielä ehtinyt kertyä. Näin arvioimme, että hän vastaisi varsin hyvin NAS:in noviisikäyttäjää.

Pilottitesti toteutettiin aivan kuin myöhemmät varsinaiset käytettävyystestitkin. Koko testi suoritettiin alku- ja loppuhaastatteluineen. Pilottitestiin kulunut aika mitattiin, ja näin pystyimme arvioimaan testin kestoa. Pilottitestin kokonaiskesto oli tasan 40 minuuttia, mikä on varsin lyhyt aika, mutta toisaalta suosituksen mukaisesti alle tunnin [Uehling 1994].

Pilottitesti paljasti muutamia puutteita testitehtävissä. Toinen harjoitustehtävä, jossa etusivun uutissyöttö piti asettaa keräämään uutisia, osoittautui liian hankalaksi. Ongelmana oli erityisesti se, että koska testikäyttäjän osioon ei oltu syötetty uutisia ei uutissyöttö myöskään näyttänyt mitään, vaikka tehtävän suoritti oikein. Tämä hämäsi luonnollisesti käyttäjää eikä ollut ollenkaan omiaan tuomaan käyttäjälle itsevarmuutta, jota testitehtävillä myös haettiin. Päätimme siis poistaa lopullisista käytettävyystesteistä tämän harjoitustehtävän.

Huomasimme myös, että testitehtävien vaatimat liitetiedostot pitää tarjota käyttäjälle cd-levyllä. Pilottitestissä tiedostot haettiin verkosta, jolloin aikaa kului aivan liikaa. Varsinaisiin käytettävyystesteihin mennessä korjattiin myös muutamia testisivuston ohjelmointivirheitä, jotka paljastuivat pilottitestissä.

5 Tulokset

5.1 Heuristisen arvioinnin tulokset

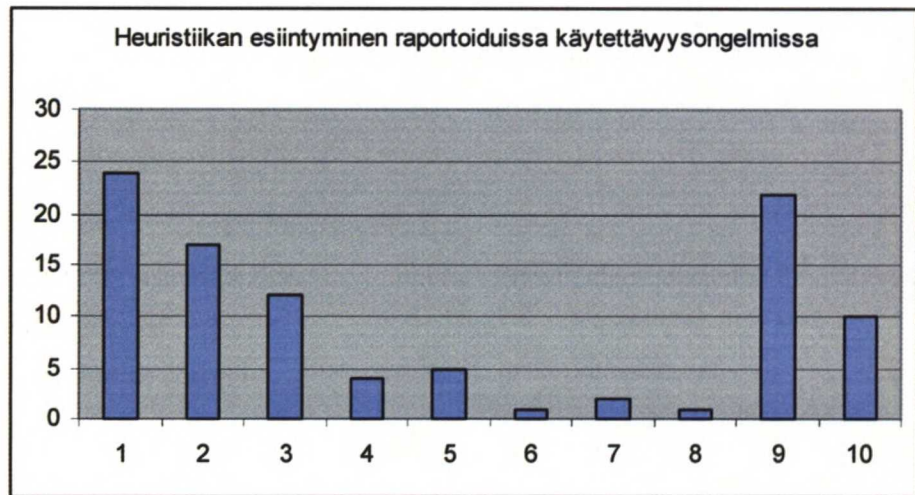
Heuristisen arvioinnin tulokset koottiin heti yksilöllisten arviointien jälkeen pidetyssä loppuyhteenvetotilaisuudessa. Arvioijien löytämät käytettävyyssongelmat käsiteltiin ryhmässä. Jokaisen käytettävyyssongelman kohdalla vakavuus arvioitiin uudelleen ja joukosta karsittiin ne, joiden ei katsottu olevan käytettävyyssongelmia eli joiden vakavuusluokan katsottiin olevan nolla. Arvioijien raportoimia samankaltaisia ongelmia yhdisteltiin päällekkäisyyksien karsimiseksi. Loppuraporttiin kirjattiin myös korjausehdotukset sekä mitä heuristiikkaan tai heuristiikkoja käytettävyyssongelma rikkoi.

Taulukko 3. Heuristisessa arvioinnissa löydetyt käytettävyyssongelmat vakavuusluokittain

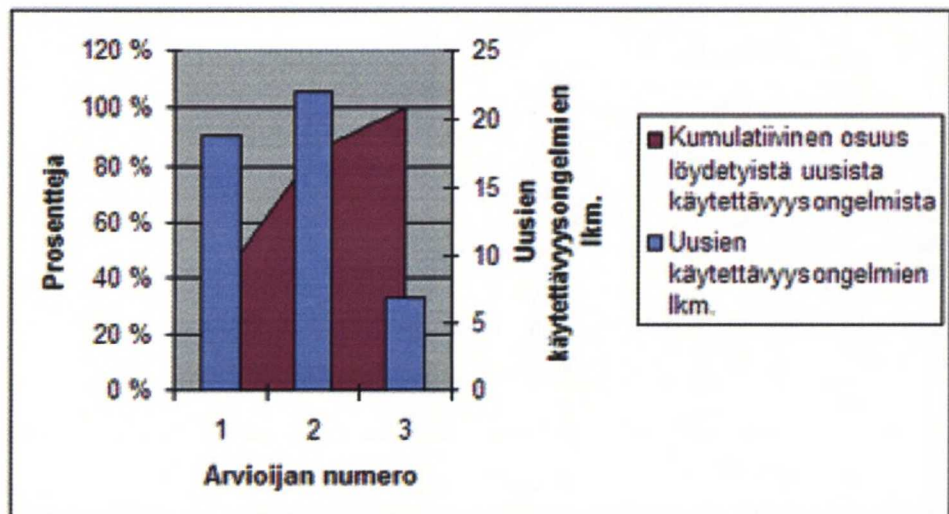
Käytettävyyssongelman vakavuus	lkm
1. Kosmeettinen	3
2. Vähäinen	24
3. Merkittävä	20
4. Katastrofi	1
Yhteensä	48

Löydetyt käytettävyyssongelma painottuivat suuresti heuristiikoille 1,2,3,9 ja 10 (Kuva 6). Osittain tämä on selitettävissä sillä, että kyseiset heuristiikat ovat niin yleisluontoisia, että useimmat käytettävyyssongelmat voidaan kirjata niihin. Arvioijien kokemus ja taidot vaikuttavat kuitenkin ratkaisevasti tuloksiin [Hertzum & Jacobsen 2001]. Tätä ongelmaa pyrittiin karsimaan käyttämällä kolmea arvioijaa, mutta jää epäselväksi olisiko arvioijia pitänyt olla enemmän.

Kuvassa 7 nähdään, että jokainen arvioija löysi runsaasti sellaisia käytettävyyssongelma, joista toiset eivät olleen raportoineet. Arvioijien lukumäärä ei siis ainakaan ollut liian suuri.



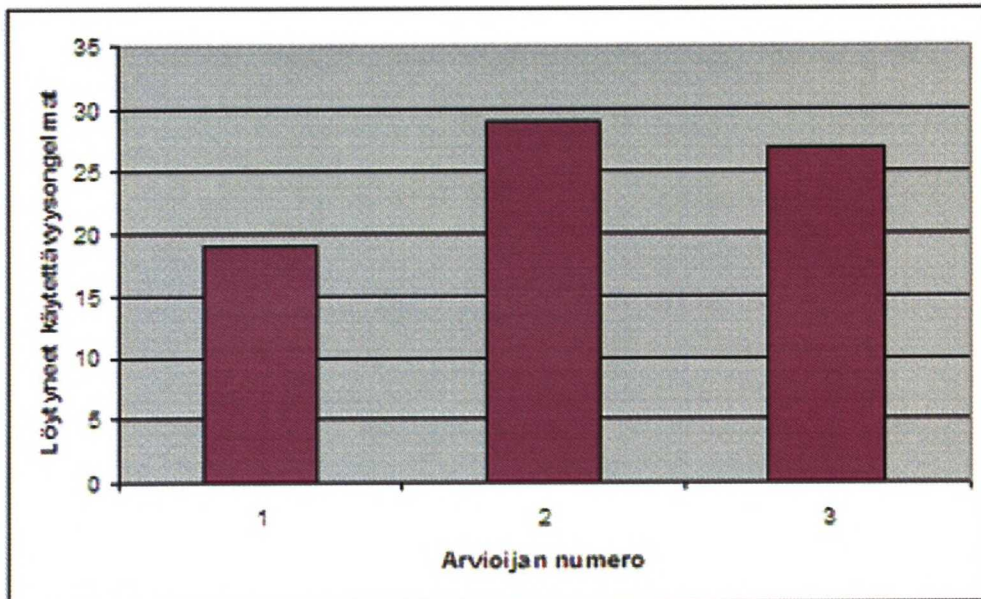
Kuva 6. Heuristisessa arvioinnissa löytyneet käytettävyyssongelmat heuristiikan mukaan jaoteltuna



Kuva 7. Heuristisessa arvioinnissa löytyneiden uusien käytettävyyssongelmien jakautuminen

5.1.1 Löydetyt käytettävyyssongelmat

Ensimmäisen tehtävän aikana löydettiin yhteensä kahdeksan käytettävyyssongelmaa. Mukana on myös koko arvioinnin ainoa vakavimman eli neljännen tason



Kuva 8. Löydetyt käytettävyyssongelmat per arvioija

käytettävyyssongelma. Tämä koski uutisen tallentumista tietokantaan, vaikka käyttäjä oli syöttänyt puutteelliset tiedot. Merkittäviksi käytettävyyssongelmiksi arvioitiin päivämääräkentät, joissa ei ohjeistettu missä muodossa syöte tulee antaa, sekä epäselvä virheilmoitus, kun päivämääräkenttään ei oltu syötetty päivämäärää. WYSIWYG-editorin toiminnassa havaittiin myös useita puutteita. Valikko esimerkiksi sisälsi toimintoja, joita uutisen lisäyksessä ei tarvita ja lisäksi editoria piti jatkuvasti siirtää muokkausalueen päältä pois.

Sivunmuokkausta käsittelevässä kakkostehtävässä löydettiin yhteensä seitsemän käytettävyyssongelmaa. Vakavimpia olivat liitetiedoston häviäminen, jos linkkitekstiä ei oltu annettu, sekä ohjeiden puuttuminen ja epäselvät termit. Kolmannessa tehtävässä arvioijat loivat uuden sivun. Sivun sijoittaminen navigaatiossa osoittautui erittäin epäselväksi ja koko toiminnon sijoitus WYSIWYG-editorissa todettiin

huonoksi. Yhteensä tehtävän aikana löydettiin viisi käytettävyyssongelmaa, jotka yhtä lukuunottamatta olivat kolmannen tason käytettävyyssongelmia.

Neljännessä tehtävässä palattiin sivunmuokkaukseen. Kuvan lisäyksestä raportoitin seitsemän käytettävyyssongelmaa, joista viisi arvioitiin vähäisiksi. Valikoista löytyi englanninkielisiä tekstejä, vaikka muuten tekstit olivat suomeksi. Lisäksi arvioijat kiinnittivät huomiota puutteelliseen ohjeistukseen sekä vaikeaselkoisiin termeihin. Viidennessä tehtävässä nousi esille sivujen siirron vaikeus. Ongelmat olivat saman suuntaisia, kuin kolmannessa tehtävässä raportoidut. Ne kirjattiin erillisiksi käytettävyyssongelmiksi, sillä aiemmin raportoidut sivun siirtoa koskevat käytettävyyssongelmat käsittelivät vain WYSIWYG-editorin toimintoja.

Kuudennessa tehtävässä arvioijien tehtävänä oli poistaa myynti-osaston alasivu. Tehtävässä huomioitiin kaksi käytettävyyssongelmaa. Seitsemännessä tehtävässä gallup-kyselyn lisäämisessä törmättiin jälleen epäselviin termeihin. Toiminnon avautuminen WYSIWYG-editorin sisällä havaittiin myös heikoksi ratkaisuksi. Kaikki kentät eivät mahtuneet kerralla ruutuun vaan esimerkiksi painikkeita sai näkyville vain vierityspalkkien avulla.

Seuraava tehtävä käsitteli salasanan vaihtoa ja merkittävimmät huomiot kirjattiin puutteellisesta palautteesta. Yhdeksäs tehtävässä kysytty Teppo Teikarisen käyttäjätunnuksen selvitys osoittautui hyvin yksinkertaiseksi, vaikkakin tunnus löytyi vain muokkaa-painiketta painamalla. Tämä kirjattiin käytettävyyssongelmaksi.

Kymmenes tehtävä oli yksi arvioinnin vaikeimmista. Työkalupakin näyttömuodon valintaan pääsi vain ”omat tiedot” -valikon kautta, mikä havaittiin epäloogiseksi. Arvioinnin viimeinen tehtävä käsitteli järjestelmästä uloskirjautumista. ”Lopeta istunto” -termin arveltiin olevan joillekin käyttäjille vieras, mutta tämäkin arvioitiin vain vähäiseksi käytettävyyssongelmaksi.

5.2 Käytettävyydestien tulokset

Käytettävyydestin alkuhaastattelussa selvitettiin käyttäjien kokemusta WWW-julkaisusta ja NAS-järjestelmän käytöstä. Vastauksien perusteella testikäyttäjät koostuivat tasaisesti noviisikäyttäjistä kokeneisiin ylläpitäjiin. Tähän viittasi muun muassa se, että pääkäyttäjinä ryhmästä toimi neljä, kun taas kolmella käyttäjällä NAS:in käytöstä ei ollut aiempaa kokemusta (Taulukko 5). Yleisesti pääkäyttäjän työtehtäviin kuuluvia käyttäjähallinnan mainitsi vastauksissaan kolme henkilöä, kun ylläpitäjän yleisimpiin työtehtäviin kuuluvan uutisten päivityksen mainitsi vastauksissaan neljä henkilöä. Muista www-julkaisujärjestelmistä käyttäjillä ei ollut juurikaan kokemusta (Taulukko 6).

Taulukko 4. Käytettävyydesteihin osallistuneiden aiempi käyttökokemus NAS-järjestelmästä

Käyttömäärä	Vastaajien lkm
Päivittäin	2
2-3 krt viikossa	2
Kerran viikossa	2
”Aika vähän”	1
NAS ei käytössä	3

Käyttömäärät vaihtelivat varsin tasaisesti päivittäisestä käytöstä viikottaiseen käyttöön (Taulukko 4). Internetin käyttö oli kaikille käyttäjille tuttua verkkopankkien ja erityisesti tiedon haun kautta. Kaikki vastaajista käyttivät työssään Microsoftin Office –ohjelmia ja muutamille myös Open Office oli tuttu. Toimisto-ohjelmien käyttökokemus näkyikin testeissä siinä, että WYSIWYG-editorin ikonit olivat kaikille tuttuja.

Taulukko 5. Vastaajien yleisimmät tehtävät NAS-järjestelmässä

Tehtävä	Vastaajien lkm
Gallupin / uutisten päivitys	4
Sisällön muutokset	5
Järjestelmän valvonta	2
Pääkäyttäjän tehtävät	4
Käyttäjähallinta	3
NAS ei käytössä	3

(Vastaaja on voinut vapaasti luetella yleisimpiä tehtäviään)

Taulukko 6. Vastaajien aiempi kokemus WWW-julkaisujärjestelmistä

Järjestelmä	Vastaajien lkm
Nedeconin/Enderon järjestelmät	3
TietoEnatorin Elevation	1
MS Sharepoint	1
Dreamweaver/Frontpage	3
HTML	3
Ei aiempaa kokemusta	2

(vastaaja on voinut luetella kaikki menetelmät, joilla on päivittänyt WWW-sivustoja)

Testitehtävät osoittautuivat ehkä liiankin helpoiksi. Käyttäjät suorittivat kaikki tehtävät lähes poikkeuksetta reilusti alle tunnissa. Nopeimmillaan koko käytettävyydestä alkua ja loppuhaastatteluihin sujui puolessa tunnissa. Kun huomioidaan, että haastatteluihin kului yleensä noin viisitoista minuuttia niin nopeimmillaan jokaisen tehtävän suoritus aika oli vain vähän reilun minuutin. On huomattava, että jo pilottitesti kesti vain 40 minuuttia, joten viitteitä testin lyhydestä oli olemassa jo ennen varsinaisten testien suorittamista. Testitehtäviä olisi siis voitu monimutkaistaa, jos pilottitestien tuloksia olisi analysoitu huolellisemmin. Varsinkin tehtävää 8 eli sanasanavaihtoa, olisi voinut jatkaa jatkotehtävällä, sillä nyt tehtävä oli liian yksinkertainen. Myös tehtävässä 7 annettiin jo tehtävänannossa turhan

selkeät ohjeet äänestysmuodon valinnasta. Tämän takia käyttäjät pystyivät suorittamaan tehtävän, vaikka koko äänestysmuodon valinnan merkitys olisi jäänyt epäselväksi. Suurempi virhe tehtiin kuitenkin siinä, että pilottitesti suoritettiin vasta heuristisen arvioinnin jälkeen. Näin ollen ei ollut enää mahdollista muuttaa varsinaisia testitehtäviä ilman, että testimenetelmien vertailtavuus olisi menetetty.

Tehtävän suorittamisen luovuttamiseen tai keskeyttämiseen johtivat lähinnä sivun siirtoon tai sijoittamiseen liittyneet tehtävät (tehtävissä 3 ja 5) sekä työkalupakin näyttömuodon valinta (tehtävä 10). Myös kuvan liittäminen aiheutti joitakin ongelmia (tehtävä 4). Muissa tehtävissä onnistuneeseen lopputulokseen pääsi erittäin suuri osa käyttäjistä (Taulukko 7).

Taulukko 7. Tehtävien onnistunut suoritus

Tehtävän nro	Onnistuminen
1	10/10
2	8/10
3	6/10
4	7/10
5	7/10
6	10/10
7	9/10
8	10/10
9	8/10
10	2/10
11	10/10

Tehtävien järjestys vaikutti onnistuneelta. Ensimmäisissä tehtävissä epäonnistumisia ei tullut kovin runsaasti, joten käyttäjien itseluottamusta ei koeteltu heti alkujännityksen vielä vallitessa. Kolmas tehtävä osoittautui yhdeksi testin vaikeimmista, joten sen olisi voinut sijoittaa hieman myöhemmäksi. Aivan testin loppuun sijoittunut todella vaikeaksi osoittautunut 10. tehtävä jäi myös selkeästi kaivelemaan muutamia käyttäjiä, sillä he halusivat heti testitehtävät suoritettuaan nähdä miten tehtävän olisi saanut onnistuneesti suoritettua. 10. tehtävä olisi ehkä pitänyt sijoittaa hieman aikaisemmaksi, jotta epäonnistumisen

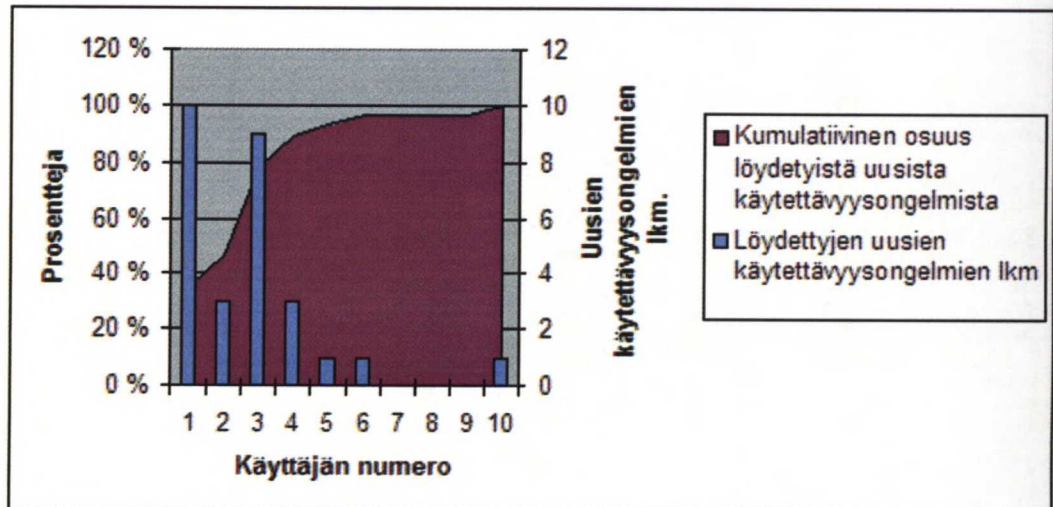
tunne ei olisi päällimmäisenä jäänyt vaivaamaan testien päätyttyä.

(Onnistuneeksi suoritukseksi ei laskettu niitä, joissa käyttäjä joutui turvautumaan käyttöohjeeseen)

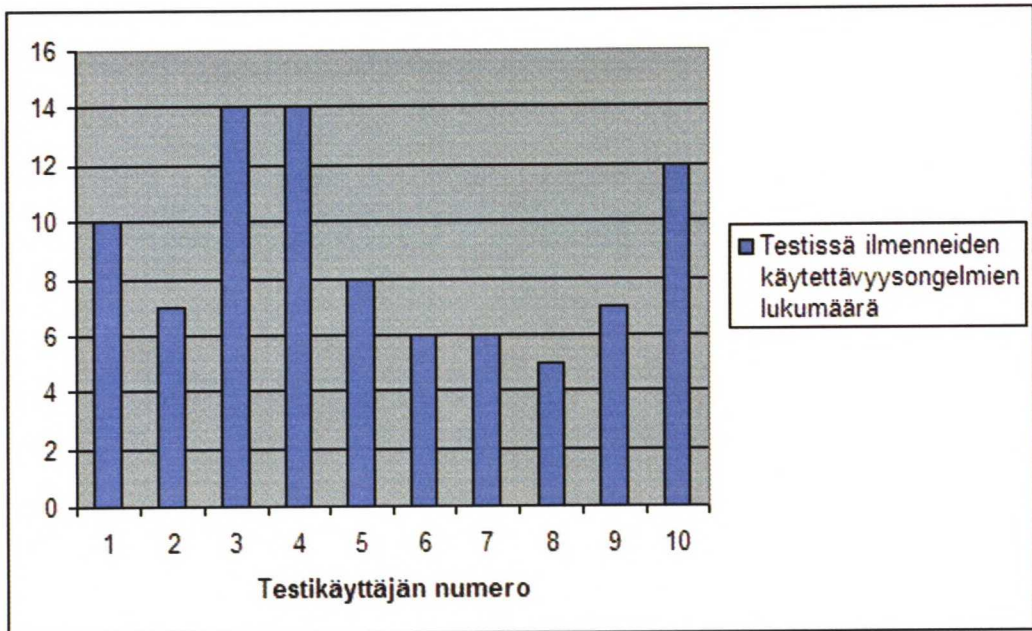
Taulukko 8. Käytettävyystesteissä ilmenneet käytettävyysongelmat vakavuusluokittain

Käytettävyysongelman vakavuus	lkm
1. Kosmeettinen	1
2. Vähäinen	12
3. Merkittävä	15
4. Katastrofi	0
Yhteensä	28

Käytettävyystesteihin osallistui kymmenen testikäyttäjää. Määrä osoittautui turhan suureksi, sillä neljännen käyttäjän jälkeen loppujen kuuden käyttäjän avulla löydettiin yhteensä vain kolme käytettävyysongelmaa, joita ei oltu aiemmissa testeissä jo löydetty (Kuva 9). Näin ollen siis luvussa 4.4.1 tehdyt perustelut yli viiden testikäyttäjän käyttämisestä osoittautuivat virheellisiksi ja testikäyttäjien joukko olisi voinut olla puolet pienempi.



Kuva 9. Käytettävyystestissä ilmenneiden uusien käytettävyysongelmien jakautuminen



Kuva 10. Käytettävyystesteissä ilmenneet käytettävyyssongelmat

5.2.1 Löydetyt käytettävyyssongelmat

Käytettävyystesteissä ilmenneet ongelmat kerättiin testien aikana tehdyistä muistiinpanoista, sekä testitilanteiden videotallenteista. Ensimmäisessä tehtävässä havaittiin yhteensä kuusi käytettävyyssongelmaa. Käyttäjät olivat epävarmoja missä muodossa päivämääräkenttiin päiväys tuli syöttää ja olivatko kentät pakollisia. WYSIWYG-editorin koko häiritsi myös useita käyttäjiä. Toisessa tehtävässä hankaluuksia aiheutti työkaluvalikon sijainti ja ikonien vaikeaselkoisuus. Tehtävän keskeytyksiä aiheutti kuvan lisäyksessä linkkitekstikentän jättäminen tyhjäksi. Tällöin WYSIWYG-editoriin ilmestynvä virheilmoitus vei kuvan liittämispainikkeen vierityspalkkien taakse piiloon. Kaksi käyttäjää ei tämän takia saanut kuvaa tallennettua ja kaksi muuta huomasi painikkeen vasta aloitettuaan tehtävän suorituksen uudelleen alusta.

Kolmannen tehtävän sai suoritettua vain kuusi kymmenestä käyttäjästä. Sivun sijoitus tiettyyn paikkaan navigaatioissa osoittautui erittäin vaikeaksi. WYSIWYG-editorissa

toiminto oli sijoitettu sellaiseen paikkaan, että vain yksi käyttäjä löysi sen. Muut tallensivat sivun ensin ja yrittivät siirtoa vasta sen jälkeen. Tämäkään ei onnistunut vaivattomasti, sillä järjestelmä ei ohjeistanut käyttäjää toiminnon kaksivaiheisesta leikkaa-liitä-periaatteesta.

Neljännessä tehtävässä tehtävän keskeytyksiä aiheutti kuvan liittäminen. Muutamat käyttäjät hämääntyivät, kun liitä-painike ei sijoittanutkaan kuvaa sivulle vaan avasi valikon, josta kuva piti vielä raahata sivulle. Esikatselutilassa kuvan kokoa pääsi myös muokkaamaan, mutta alkuperäistä kokoa ei saanut palautettua mitenkään. Tosin koon muuttaminen muokkaustilassa ei lopulta vaikuttanut kuvan lopulliseen kokoon, sillä sivulle siirron jälkeen kuva näkyi taas alkuperäisessä koossaan.

Viidennessä tehtävässä aiemmin paljastuneet ongelmat sivun siirrossa uusiutuivat. Ainoana uutena käytettävyysongelmana huomattiin käyttäjien joutuvan toistuvasti pienentämään työkaluvalikkoa pois navigaation päältä. Kuudennen tehtävän saivat kaikki kymmenen käyttäjää suoritettua onnistuneesti. Sivun poistoa hankaloitti, että sivun sai poistettua vain navigaatioissa ylemmällä tasolla olevan sivun kautta.

Gallupin liittäminen sivulle oli käyttäjien tehtävänä seitsemännessä tehtävässä. Vaikka toiminto sisälsi epäselviä termejä ja käyttäjän ohjaaminen oli puutteellista, vain yksi käyttäjä ei saanut tehtävää suoritettua. Kahdeksas tehtävä osoittautui todella helpoksi eikä siinä havaittu käytettävyysongelmia.

Yhdeksännessä tehtävässä ongelmia aiheutti se, että etsitty käyttäjätunnus löytyi vain menemällä käyttäjätiedoissa muokkaustilaan. Lähes kaikki käyttäjät kamppailivat tämän ongelman kanssa ja se aiheutti myös kahdelle käyttäjälle tehtävän keskeytyksen. Kymmenennessä tehtävässä työkaluvalikon näyttömuodon valinta oli sijoitettu niin epäloogiseen paikkaan, että vain kaksi käyttäjää onnistui tehtävän suorituksessa. Heistäkin toinen löysi valikon vasta pitkän etsimisen jälkeen.

Viimeisenä tehtävä oli järjestelmästä ulos kirjautuminen. Tämä ei aiheuttanut kellekään minkäänlaisia ongelmia.

5.3 Arviointimenetelmien tulosten vertailu

5.3.1 Tulosten vertailu tehtävittäin

Ensimmäisessä tehtävässä päivämääräkenttien syöttömuodon ohjeistukset puutteet ja tallennuspainikkeen huono sijainti havaittiin molemmilla menetelmillä. Heuristisella arvioinnilla havaittiin koko tutkimuksen ainoa neljännen tason käytettävyysongelma, joka löytyi, kun päivämääräkenttiin syötettiin tahallaan päivämääräksi sopimaton teksti.

Toisessa tehtävässä käytettävyystesteillä ei löydetty yhtään sellaista ongelmaa, mitä ei olisi heuristisella arvioinnilla löydetty. Heuristisella arvioinnilla löydettiin näiden lisäksi myös heikkoon ohjeistukseen liittyviä käytettävyysongelmia. Kolmannessa tehtävässä menetelmillä havaittiin samoja sivun siirtoon liittyviä ongelmia. Heuristisella arvioinnilla löydettiin lisäksi kaksi ja käytettävyystesteillä yksi ongelma, jota toisella menetelmällä ei löydetty.

Neljännessä tehtävässä käytettävyystesteillä löydettiin vain kaksi ongelmaa, jotka oli jo havaittu heuristisella arvioinnilla. Heuristisella arvioinnilla löydettiin näiden lisäksi runsaasti epäselviin termeihin ja heikkoon ohjeistukseen liittyviä ongelmia. Viidennessä tehtävässä käytettävyystesteillä havaittiin käyttäjien joutuvan toistuvasti pienentämään työkaluvalikko navigaation päältä pois. Käytettävyystesteillä tämän tyylisten ongelmien havainnointi oli helpompaa, sillä arvioijat pääsivät näkemään usean käyttäjän kamppailemassa virheen parissa.

Kuudennessa tehtävässä molemmilla menetelmillä huomattiin, ettei sivun poisto onnistunut kuin navigaatioissa sivua ylemmältä tasolta. Käytettävyystesteissä muutamia käyttäjiä hämäsi sivun poistamisesta tuleva ilmoitus ”sivu siirretty

roskakoriin” ja he etsivät vielä tämän jälkeen mahdollisuutta roskakorin tyhjentämiselle. Tätä ei oltu huomioitu testejä suunniteltaessa.

Gallupin lisäämisestä raportoitiin molemmissa menetelmässä runsaasti käytettävyyssongelmia. Niistä vain kaksi löydettiin kummallakin menetelmällä. Heuristisessa arvioinnissa arvioijat kokeilivat saada aikaan virhetilanteita ja näin löydettiin yksi käytettävyyssongelma, jota ei käytettävyystesteissä ilmennyt.

Käytettävyystesteillä ei kahdeksannessa tehtävässä löydetty yhtäkään käytettävyyssongelmaa. Heuristisella arvioinnilla tehtävässä raportoitiin ongelmasta, joka ilmeni vain jos järjestelmässä kytkettäisiin automaattinen kirjautuminen päälle. Tällaista ongelmaa ei olisi käytettävyystesteillä voitu löytää, sillä kyseinen toiminto ei ollut testijärjestelmässä päällä. Heuristisessa arvioinnissakin tämän ongelman löytämiseksi tarvittiin henkilö, jolla oli vankka tuntemus järjestelmän rakenteesta ja toiminnasta testijärjestelmää poikkeavilla asetuksilla.

Työkaluvalikon näyttömuodon valinnan epälooginen paikka ”omat tiedot” –valikossa huomattiin molemmilla menetelmillä. Heuristisella arvioinnissa arvioijat kokeilivat myös tallentuuko valittu näyttömuoto niin, että se on käytettävissä kun käyttäjä kirjautuu järjestelmään uudestaan. Tämä ei kuulunut testitehtäviin eikä tähän liittynyttä käytettävyyssongelmaa löytynyt käytettävyystesteillä, koska testikäyttäjät tekivät vain mitä testitehtävässä pyydettiin.

Viimeisessä tehtävässä käytettävyystesteillä ei löydetty käytettävyyssongelmia. Heuristisessa arvioinnissa arvioijat ehdottivat ”lopeta istunto” –termin selkeyttämistä, epäillen, että se voi joillekin peruskäyttäjille olla vieras.

5.3.2 Vertailun yhteenveto

Heuristisessa arvioinnissa löytyi selvästi enemmän käytettävyyssongelmia kuin käytettävyystesteissä (Taulukko 9). Näiltä osin tulokset vastasivat myös aiempia

tutkimuksia (Jeffries et al. 1991, Karat et al. 1992, Nielsen 1994b). Käytettävyyystesteillä löytyi suhteellisesti enemmän merkittäviä (vakavuusluokka 3) käytettävyyssongelmia verrattuna heuristiseen arviointiin. Tämän Nielsenkin raportoi omassa tutkimuksessaan [Nielsen 1994b]. Toisaalta taas käytettävyysskatastrofeja ei käytettävyyystesteillä löytynyt yhtäkään, kun taas heuristisella arvioinnilla löytyi yksi.

Taulukko 9. Löydetty käytettävyyssongelmat yhteensä (sisältää ongelmat, jotka löydettiin kummallakin menetelmällä)

Käytettävyyssongelman vakavuus	HA*	KT**	Yht (***)
1. Kosmeettinen	2	1	3
2. Vähäinen	22	12	26
3. Merkittävä	23	15	26
4. Katastrofi	1	0	1
Yhteensä	48	28	56

* HA = Heuristinen arviointi

** KT = Käytettävyystestit

*** Ei laskettu käytettävyyssongelmien kaksoiskappaleita, joissa sama ongelma löytynyt jo heuristisella arvioinnilla

Merkittävä ero heuristisen arvioinnin ja käytettävyystestien tuloksia vertailtaessa on sen sijaan vertailu siitä, kuinka suuri osa menetelmillä löydettyistä ongelmista löytyi vain toisella menetelmällä (Taulukot 10 ja 11). Heuristinen arviointi osoittautui tässä vertailussa tehokkaammaksi, sillä erityisesti vakavuusluokaltaan merkittäviksi (3) ja katastrofiksi (4) arvioiduista ongelmista huomattava osa löytyi vain heuristisella arvioinnilla. Kaikista käytettävyyssongelmista (heuristisessa arvioinnissa ja käytettävyyystesteissä löytyneet yhteensä) puolet löytyi vain heuristisella arvioinnilla. Käytettävyyystesteissä löytyneistä kaikista käytettävyyssongelmista 29% ei löydetty heuristisella arvioinnilla. Näistä yli puolet kuului vakavuusluokiltaan alhaisiin eli kosmeettisiin (1) tai vähäisiin (2).

Taulukko 10. Käytettävyysongelmat, jotka löydettiin vain heuristisella arvioinnilla

Vakavuusluokka	lkm	% heuristisella arvioinnilla löydettyistä kyseisen luokan käytettävyysongelmissa	% kyseisen luokan käytettävyysongelmiä kokonaismäärästä
1. Kosmeettinen	2	100%	67%
2. Vähäinen	14	64%	53%
3. Merkittävä	11	48%	42%
4. Katastrofi	1	100%	100%
Yhteensä	28	58%	50%

Taulukko 11. Käytettävyysongelmat, jotka löydettiin vain käytettävyystesteillä

Vakavuusluokka	lkm	% käytettävyystesteillä löydettyistä kyseisen luokan käytettävyysongelmissa	% kyseisen luokan käytettävyysongelmiä kokonaismäärästä
1. Kosmeettinen	1	100%	33%
2. Vähäinen	4	33%	15%
3. Merkittävä	3	20%	12%
4. Katastrofi	0	0%	0%
Yhteensä	8	29%	14%

Heuristisessa arvioinnissa esiin tuli myös muutama selvä ohjelmistovirhe, joita käytettävyystesteissä ei ilmennyt. Yksi syy näihin oli se, että heuristisessa arvioinnissa arvioijat kokeilivat tahallaan syöttää syötekenttiin virheellisiä arvoja. Käytettävyystesteissä tällaisia virhetilojen metsästämisestä ei esiintynyt ja nämä ohjelmistovirheet jäivät löytymättä.

Käytettävyydestien selkeänä etuna oli arvioijien mahdollisuus päästä seuraamaan järjestelmän erilaisia käyttötapoja. Testeistä kirjattiinkin käytettävyysoongelmiksi käyttöä hidastavia toimintoja, kuten esimerkiksi se, että työkaluvalikko täytyi toistuvasti pienentää navigaation päältä. Tällaisten käytettävyyso ongelmien löytäminen heuristisella arvioinnilla oli selkeästi hankalampaa, sillä eri arvioijat eivät päässeet seuraamaan toisten arvioijien suoritusta, joten ongelmien toistuminen paljastui vasta, jos toinenkin arvioija oli raportoinut saman ongelman. Käytettävyydesteissä ongelman vakavuuden arviointi oli tältä osin helpompaa, kun testien jälkeen voitiin selkeästi katsoa kuinka monessa testissä ongelma oli esiintynyt ja miten se oli vaikuttanut tehtävän suorittamiseen. Heuristisessa arvioinnissa vakavuuden arviointi oli hankalampaa, sillä ei ollut varmuutta kuinka paljon käytettävyyso ngelma vaikuttaa todellisiin käyttäjiin.

Arviointimenetelmien vaatiman ajan vertailun osalta tehtiin virheitä, joiden takia tarkka vertailu on mahdotonta. Heuristisesta arvioinnista mitattiin vain itse arviointiin ja loppuyhteenvetoon käytetty aika. Käytettävyydesteistä kirjattiin vain testiin käytetty aika, mutta esimerkiksi testipaikoille matkustamiseen kulunutta aikaa ei otettu huomioon. Myöskään itse testien valmisteluihin, kuten testikäyttäjien hankkimiseen käytettyä aikaa ei mitattu. Tästä huolimatta voidaan kuitenkin sanoa, että heuristinen arviointi on huomattavasti nopeampi ja kevyempi toteuttaa, vaikka käytetty aika arvioitaisiin vain päivän tarkkuudella. Heuristinen arviointi suoritettiin päivässä, jolloin valmistui myös loppuyhteenveto tuloksista. Käytettävyydesteissä pelkät testit kestivät kaksi viikkoa, tosin itse testit veivät päivistä vain puolet. Nopeimmillaan testit olisi voitu suorittaa siis viikossa, mutta tähän ei vielä sisälly tulosten analysointia. Videonauhoitusten läpikäynti oli erittäin hidasta ja siihen kului vielä useita työpäiviä. Heuristisella arvioinnilla olisi pystytty huomioimaan eri käyttöroolien tarpeet tekemällä arvioinnit eri rooleille suunnatuilla tehtävillä. Se on sisällönhallintajärjestelmää tutkittaessa erittäin tärkeää. Käytettävyydesteissä eri

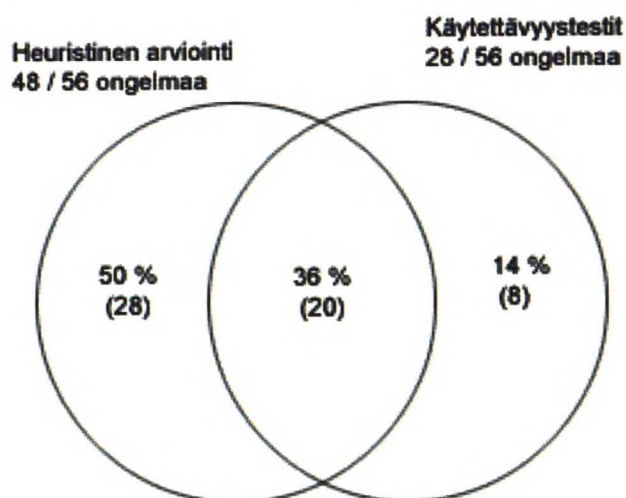
roolien huomioiminen olisi vaatinut uusien, näihin käyttörooleihin kuuluvien, testikäyttäjien osallistumista, joka olisi entisestään lisännyt testauksen raskautta.

Käytettävyystesteissä löytyneiden käytettävyyso Ongelmien korjausehdotuksien laatimista helpotti se, että testien loppuhaastattelussa käyttäjät pystyivät myös itse antamaan parannusehdotuksia. Tosin näihin ehdotuksiin tulee suhtautua varsin kriittisesti sillä kyseessä oli kuitenkin subjektiivinen mielipide [Nielsen 1993]. Arviointityötä auttoi myös se, että käytettävyystesteissä löytyneet ongelmia voitiin pitää hyvinkin todellisina ongelmina loppukäyttäjille, kun taas heuristisessa arvioinnissa virheellisten hälytysten mahdollisuus piti pitää mielessä jokaista löytynyttä ongelmaa arvioitaessa.

6 Analyysi

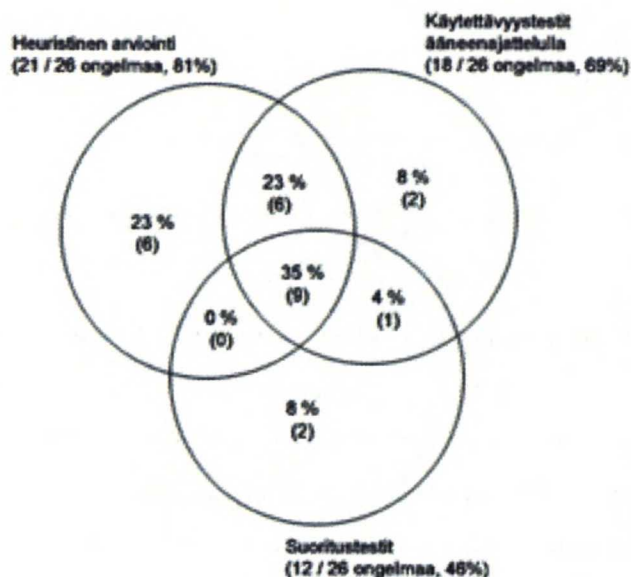
6.1 Tulosten vastaavuus aiempiin tutkimuksiin

Useissa aiemmissä käytettävyystestien ja heuristisen arvioinnin vertailuissa johtopäätöksenä on ollut, että menetelmät ovat enemmänkin toisiaan täydentäviä kuin korvaavia [Jeffries et al. 1991, Karat et al. 1992, Nielsen 1994b]. Tutkimuksissa raportoitiin menetelmien löytävän suureksi osaksi eri käytettävyyso ongelmia.



Kuva 11. Löydettyjen käytettävyyso ngelmien jakautuminen eri menetelmien välillä nyt tehdyssä tutkimuksessa

Tässä työssä saadut tulokset eivät ole samassa linjassa. Aiemmissä tutkimuksissa heuristisella arvioinnilla löydettiin paljon ongelmia, joita käytettävyyso ntesteillä ei löydetty, mutta käytettävyyso ntesteillä löydettiin lähinnä ongelmakohtia, jotka oli jo löydetty heuristisella arvioinnilla. Vain 14 prosenttia kaikista raportoiduista käytettävyyso ngelmista löydettiin pelkästään käytettävyyso ntesteissä.



Kuva 12. Virzin tutkimusryhmän [Virzi et al. 1993] saamat tulokset arviointimenetelmien vertailussa

Hieman tämän tutkimuksen kanssa saman suuntaisia tuloksia saatiin Virzin tutkimusryhmän tekemässä kolmen arviointimenetelmän vertailussa [Virzi et al. 1993]. Siinä käytettävyyshälytyksellä löydettiin lähes samassa suhteessa sellaisia ongelmia, joita ei heuristisella arvioinnilla löydetty. Toisaalta heidän tutkimuksessaan yli puolet käytettävyyshälytyksistä löydettiin sekä heuristisella arvioinnilla, että käytettävyyshälytyksellä, mikä ei taas ole yhteneväistä tässä tutkimuksessa saatujen tulosten kanssa.

On kuitenkin huomioitava, että kummassakin tutkimuksessa kerätty aineisto on niin suppea, ettei siitä voida tehdä tilastollisesti luotettavia päätelmiä. Virzin työryhmän tutkimuksessa löydettyjen käytettävyyshälytysten kokonaismäärä on vain 26 ja nyt tehdyssä diplomityössä aineisto on vain 56 käytettävyyshälytysten suuruinen. Heuristisessa arvioinnissa tulokset pohjautuvat tämän diplomityön osalta vain kolmen arvioijan löydöksiin. Käytettävyyshälytyksissä testikäyttäjiä oli kymmenen, mutta tuloksia analysoi vain kaksi testin järjestäjää. Näin ollen yli kymmenenkin prosentin

virhemarginaali on täysin mahdollinen ja erot Virzin tutkimusryhmän tuloksiin mahtuvat tämän marginaalin sisälle.

6.1.1 Virheelliset hälytykset

Heuristisen arvioinnin osalta epäilyksiä aiheuttavat myös virheellisten hälytysten mahdollisuus. Gray ja Salzman [Gray & Salzman 1998] mainitsevat heuristisen arvioinnin heikkoudeksi lukuisat käytettävyysoingelmat, joita ei muilla arviointimenetelmillä pystytä vahvistamaan. Tässäkin tutkimuksessa heuristisella arvioinnilla löydettiin huomattava määrä käytettävyysoingelmia, joita käytettävyyssysteissä ei löytynyt. Näistä virheellisten hälytysten karsiminen riittäväällä luotettavuudella ei nykyisellä aineistolla ole mahdollista. Tarvittaisiin ainakin kolmannella arviointimenetelmällä suoritettut testit, jonka tuloksiin vertailemalla osa heuristisella arvioinnilla löydetyistä käytettävyysoingelmista voitaisiin mahdollisesti poistaa epäiltyjen listalta. Tosin osa käytettävyysoingelmista, joita ei käytettävyyssysteillä löydetty, voidaan kuitenkin varmuudella pitää todellisina. Näihin kuuluvat mm. selvät ohjelmistovirheet.

6.1.2 Arvioinnin puutteet

Luotettavampia johtopäätöksiä voitaisiin tehdä, jos heuristisia arviointeja olisi tehnyt useampi rinnakkainen arviointiryhmä toisistaan riippumatta. Myös käytettävyyssysteissä olisi tällöin pitänyt käyttää useita rinnakkaisia testausryhmiä, joilla jokaisella olisi ollut omat testikäyttäjryhmänsä. Tällöin ongelmaksi olisi tullut eri testikäyttäjryhmien väliset erot esimerkiksi käyttökokemuksen ja taitojen osalta.

Käytettävyyssysteien osalta voidaan kuvan 8 perusteella sanoa, että testikäyttäjien lukumäärä sinänsä oli täysin riittävä, jopa turhan suuri. Neljännen testikäyttäjän jälkeen ei juurikaan löydetty enää uusia käytettävyysoingelmia. Näin ollen voidaan sanoa, että Gray ja Salzmanin [Gray & Salzman 1998] raporttoiman villikortti-ilmiön vaikutus on tuloksissa minimoitu, eikä siis yksittäisen käyttäjän keskimääräistä huomattavasti parempi tai huonompi suoritus vaikuttanut tuloksiin.

Tulosten analysoijien kokemus oli sen sijaan osittain puutteellinen. Arvioijien kokemus käytettävyydestien järjestämisestä ja heuristisen arvioinnin suorittamisesta keskittyi lähinnä käynnissä olleiden opintojen harjoitustöihin. Näin ollen ketään analysoijista ei voida pitää erityisen kokeneena käytettävyyssiantuntijana. Monia käytettävyysoongelmia jäi todennäköisesti huomaamatta kummankin menetelmän osalta. Myös testitehtävien laadinnassa oli puutteita kuten luvussa 5.2 mainittiin.

Nyt saatujen tulosten luotettavuutta kyseenalaistaa myös se, että käytettävyydestien tulosten analysoinnista vastasivat samat henkilöt, jotka osallistuivat myös heuristiseen arviointiin. Heuristinen arviointi tulokset olivat käytettävyydestien tuloksia analysoitaessa jo tiedossa. Tämä saattoi vaikuttaa siihen mitä seikkoja kirjattiin käytettävyydesteistä käytettävyysoongelmiksi. Tietoisesti tähän ei kuitenkaan pyritty ja on huomioitava, että aikaa arvioinnin ja ensimmäisen käytettävyydestin välissä ehti kulua lähes kuukausi.

6.1.3 Testien puutteet

Nyt tehdyissä käytettävyydesteissä ja heuristisessa arvioinnissa keskityttiin mittaamaan käytettävyyso ongelmien määrää ja laatua. Niissä ei kuitenkaan huomioitu mitatun järjestelmän soveltuvuutta tehtäväänsä. Öörni [Öörni 2003] mainitsee, että ennen testauksen suoritusta olisi pitänyt asettaa tavoitetasot, esimerkiksi tehtävän suoritukseen kulunut aika, johon tuloksia olisi verrattu. Näin saaduista mittaustuloksia olisi voitu tehdä arvioita järjestelmän soveltuvuudesta tehtäväänsä. Tässä diplomityössä päätavoitteena oli kuitenkin kahden arviointimenetelmän vertailu ja tällaisten tavoitetasojen käyttäminen vertailukelpoisesti sekä heuristisessa arvioinnissa, että käytettävyydesteissä ei ole mahdollista.

Testijärjestelyiden ongelmana oli erityisesti se, että heuristinen arviointi suoritettiin ennen käytettävyydestien pilottitestiä. Näin pilottitestin hyödyt menetettiin, kun varsinaisia testitehtäviä ei enää voitu muuttaa, ilman, että heuristinen arviointi olisi

pitänyt uusia. Toisaalta, kun samat arvioijat olivat mukana sekä heuristisessa arvioinnissa, että käytettävyystesteissä, olisi pilottitestin suorittaminen ennen heuristista arviointia todennäköisesti vaikuttanut heuristisen arvioinnin tuloksiin. Pilottitestissä ilmenneet käytettävyysoongelmat olisivat muistuneet arvioijien mieleen heuristisessa arvioinnissa ja tällöin tuhonnut arviointimenetelmien riippumattomuuden toisistaan.

7 Johtopäätökset ja yhteenveto

Tässä diplomityössä vertailtiin kahden käytettävyyden arviointimenetelmän soveltuvuutta sisällönhallintajärjestelmän käytettävyyden arviointiin. Kiinnostuksen kohteena oli selvittää, onko tutkitun järjestelmän käyttö muista www-järjestelmistä niin poikkeavaa etteivät aiemmat näiden arviointimenetelmien vertailututkimuksien tulokset päde tähän järjestelmään.

Teoreettisessa osuudessa (luvut 2 ja 3) esiteltiin vertailtavina olleet käytettävyyden arviointimenetelmät ja tehtiin yleinen johdanto sisällönhallintajärjestelmien toimintaan ja rakenteeseen. Esimerkkitapauksena luotiin katsaus NAS-sisällönhallintajärjestelmään. Työn loppuosassa suunniteltiin ja toteutettiin käytettävyyden arvioinnit aiemmin esitellyillä arviointimenetelmillä.

7.1 Vastaus tutkimuskysymykseen

Nyt saadut tulokset kannustavat selkeästi heuristisen arvioinnin käyttöön käytettävyydestien sijaan. Menetelmä soveltui hyvin sisällönhallintajärjestelmän käytettävyyden arviointiin, sillä arvioinneissa voitiin huomioida järjestelmän käyttäjien useat eri käyttöroolit ja suunnata arviointi erityisesti tiettyjen roolien tarpeita huomioimaan. Heuristisella arvioinnilla löydettiin lähes samat käytettävyysongelmat kuin käytettävyydesteilläkin. Menetelmä vaati selkeästi vähemmän aikaa ja tulosten analysointi voitiin suorittaa samana päivänä arviointien jälkeen. Tässä tutkimuksessa tosin käytettävyydestit suoritettiin liiankin raskaalla tavalla, sillä testikäyttäjien määrän olisi voinut pudottaa puoleen tulosten juurikaan kärsimättä.

Käytettävyydestien käyttö on edelleen paikallaan heuristisen arvioinnin tuloksia vahvistamassa. Käytettävyydestien tuloksia pidetään edelleen vakuuttavimpina, kun testeistä voidaan tuottaa kvantitatiivista aineistoa esimerkiksi suoritusajoista. Lisäksi päättäjät voidaan tuoda tarkkailemaan testejä tai jopa osallistumaan niihin.

Käytettävyyshäestit soveltuvat hyvin suurempiin asiakasprojekteihin, joissa asiakas erikseen vaatii tuotteen käytettävyyden mittaamista ja on myös valmis maksamaan siitä.

Heuristinen arviointi taas soveltuu keveytensä ansiosta myös aivan pieniinkin kehitysprojekteihin ja tarvittaessa sitä voidaan hyödyntää useaan kertaan projektin eri vaiheissa. Kevyimmillään se toimii ilman ennalta valmisteltuja tehtäviä. Tällöin muutama käytettävyyden arvioija käyttää puolesta tunnista tuntiin sovelluksen läpikäyntiin kirjaten ongelmakohdat esimerkiksi Post-it –lapuille ja ongelmakohdat voidaan läpikäydä sovelluskehittäjien kanssa heti arvioinnin perään. Koko testaus voidaan näin suorittaa jopa tunnissa, mikä ei pitäisi olla este edes pienimmissä projekteissa. Heuristisella arvioinnilla voidaan selvästi käytettävyyshäestejä vaivattomammin tehdä useita arviointeja, joissa jokaisessa keskitytään tietyn käyttöroolin tarpeisiin. Tämä on erityisen tärkeää sisällönhallintajärjestelmän käytettävyyttä arvioitaessa.

7.2 Jatkotutkimuksen aiheita

Tässä tutkimuksessa heuristinen arviointi suoritettiin Nielsenin ja Molichin heuristiikalla. Tämä heuristiikka on kehitetty varsin yleiskäyttöiseksi, jotta sitä voidaan hyödyntää ohjelmistotyypistä riippumatta. Selvittämättä jäi olisiko heuristisen arvioinnin tehokkuutta voitu entisestään parantaa kehittämällä sisällönhallintajärjestelmän erityispiirteet huomioiva heuristiikka.

Heuristisesta arvioinnista on myös kehitetty runsaasti muunnoksia, kuten esimerkiksi osallistuva heuristinen arviointi [Muller & McClard 1995, Muller et al. 1998], joka yhdistelee käytettävyyshäestien ja heuristisen arvioinnin parhaimpia ominaisuuksia. Menetelmä voisi tarjota ratkaisun niihin puutteisiin, jotka tässä tutkimuksessa heuristisessa arvioinnissa havaittiin käytettävyyshäesteihin verrattuna.

8 Lähteet

Brooks, P. 1994. Adding value to usability testing. Kirjasta: J. Nielsen and R.L. Mack (eds.), Usability Inspection Methods. John Wiley & Sons. s. 255- 271. ISBN: 0-471-01877-5

Carroll, J. & Rossom, M. 1992. Getting around the task-artifact cycle: how to make claims and design by scenario. ACM Transactions on Information Systems (TOIS), Vol. 10 , nro 2 (April 1992). s. 181 -212. ACM Press. ISSN:1046-8188

Cockton, G., & Woolrych, A. 2002. Sale must end: Should discount methods be cleared off HCI's shelves? Interactions, Vol. 9, nro 5, s. 13-18. ACM Press. ISSN:1072-5220

Connell, I.W. 2000. The Use of Cognitive and Other Principles in Usability Evaluation: principles versus heuristics and cumulative problem curves. Doctoral thesis, University of York Department of Psychology, October 2000. [Viitattu 24.10.2004] Saatavilla: <http://www.cs.ucl.ac.uk/staff/i.connell/Thesispage.html>

Cordes, R.E. 2001. Task-selection bias: A case for user-defined tasks. International Journal of Human-Computer Interaction Vol. 13, nro 4, s. 411-419. Lawrence Erlbaum Associates, Inc. ISSN: 1044-7318

Dennis, A. 1998. Lessons from three years of Web development. Communications of the ACM, Vol. 41, nro 7, s. 112-113. ACM Press. ISSN: 0001-0782

Dicks, R.S. 2002. Mis-usability: on the uses and misuses of usability testing. Proceedings 20th annual international conference on Computer Documentation, s. 26-30. ACM Press. ISBN: 1-58113-543-2

Dray, S., Siegel, D. 1999. Penny-wise, pound-wise: making intelligent trade-offs in planning usability studies. *Interactions*, Vol. 6, nro 3, s. 25-30. ACM Press. ISSN: 1072-5220

Gould, J. D. and Lewis, C. 1983. Designing for usability: Key principles and what designers think. Proceedings of the SIGCHI conference on Human Factors in Computing Systems. s. 50-53. ACM Press. ISBN:0-89791-121-0

Gray, W. D., & Salzman, M. C. 1998. Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human-Computer Interaction*, Vol. 13, nro 3, s. 203-261.

Hartson, H.R., Andre, T.S. & Williges, R.C. 2003. Criteria for evaluating usability evaluation methods. *International Journal of Human-Computer Interaction*, Vol. 15, nro. 1, s. 145-181.

Hertzum, M. & Jacobsen, N.E. 2001. The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*, 13(4), s. 421-443. [Viitattu 22.10.2004] Saatavilla: http://www.akira.ruc.dk/~mhz/Research/Publ/IJHCI2001_preprint.pdf

Jacobsen, N.E., Hertzum, M. & John, B.E. 1998. The evaluator effect in usability tests. Proceedings of CHI98, s. 255-256. ACM Press.

Jeffries, R. 1994. Usability problem reports: Helping evaluations communicate effectively with developers. Kirjasta: J. Nielsen and R.L. Mack (eds.), *Usability Inspection Methods*. John Wiley & Sons. s. 273- 294. ISBN: 0-471-01877-5

Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. M. 1991. User interface evaluation in the real world: A comparison of four techniques. Proceedings of CHI'91. ACM Press.

John, B. 1995. Why GOMS? Interactions, Vol. 2, nro 4, s. 80-89. ACM Press. ISSN:1072-5520.

John, B. & Marks, S. 1997. Tracking the effectiveness of usability evaluation methods. Behaviour and Information Technology, 16, s. 188-202. [Viitattu 11.10.2004]

Saatavilla:

<http://citeseer.ist.psu.edu/cache/papers/cs/3117/http:zSzzSzreports-archive.adm.cs.cmu.edu/zSzanonzSz1996zSzCMU-CS-96-160.pdf/john94tracking.pdf>

Kamper, R. J. 2002. Extending the usability of heuristics for design and evaluation: Lead, follow, and get out the way. International Journal of Human Computer Interaction, 14(3&4), s. 447-462. [Viitattu 23.10.2004] Saatavilla: http://www.leaonline.com/doi/pdf/10.1207/S15327590IJHC143&4_10

Karat, C-M. 1994. A comparison of user interface evaluation methods. Kirjasta: J. Nielsen and R.L. Mack (eds.), Usability Inspection Methods. John Wiley & Sons. s. 203- 233. ISBN: 0-471-01877-5

Karat, C-M., Cambell, R. & Fiegel, T. 1992. Comparisons of empirical testing and walkthrough methods in user interface evaluation. Proceedings of the SIGCHI conference on Human factors in computing systems CHI'92. ACM. ISBN:0-89791-513-5

Law, L. C., & Hvannberg, E. T. 2002. Complementarity and convergence of heuristic evaluation and usability test: A case study of UNIVERSAL brokerage platform. In

Proceedings of the second Nordic conference on human-computer interaction. s. 71-80. ACM Press. ISBN:1-58113-616-1

Lewis, C., Polson, P., Wharton, C. & Rieman, J. 1990. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people. s. 235 – 242. ACM Press. ISBN:0-201-50932-6

Microsoft Corporation, 1999. Windows User Experience. Microsoft professional editions. Microsoft press. ISBN 0-73-560566-1

Molich, R. & Nielsen, J. 1990. Improving human-computer dialogue: What designers know about traditional interface design. Communications of ACM, Vol. 33, nro 3. s. 338 – 348. ACM Press.

Molich, R., Bevan, N., Curson, I., Butler, S., Kindlund, E., Miller, D. & Kirakowski, J. 1998. Comparative Evaluation of Usability Tests. Proceedings of the Usability Professionals Association 1998 (UPA98) Conference. [Viitattu 22.10.2004] Saatavilla: <http://www.dialogdesign.dk/tekster/cue1/cue1paper.doc>

Muller, M. & McClard, A., 1995. Validating an extension to participatory heuristic evaluation: quality of work and quality of work life. Conference companion on Human factors in computing systems. s. 114-115. ACM Press. ISBN:0-89791-755-3.

Muller, M., Matheson, L., Page, C. & Gallup, R., 1998 Participatory heuristic evaluation. Interactions, Vol. 5, nro 5, s. 13-18. ACM Press. ISSN:1072-5520

Nielsen, J. & Molich, R. 1990. Heuristic evaluation of user interfaces. Proceedings of CHI 1990. s. 249-256. ACM Press.

Nielsen, J. 1992. Finding Usability Problems Through Heuristic Evaluation. CHI '92 Conference Proceedings on Human Factors in Computing. s. 373-380. ACM Press. ISBN 0-89791-513-5.

Nielsen, J. 1993. Usability Engineering. Academic Press. ISBN 0-12-518495-0

Nielsen, J. 1994. Enhancing the Explanatory Power of Usability Heuristics. Proceedings of the SIGCHI conference of Human Factors in Computing Systems. s. 152 – 158. ACM Press.

Nielsen, J. 1994b. Heuristic evaluation. Kirjasta: J. Nielsen and R.L. Mack (eds.), Usability Inspection Methods. John Wiley & Sons. s. 25- 62. ISBN: 0-471-01877-5

Nielsen, J. 1995. Usability Inspection Methods. CHI '95 Conference Companion on Human Factors in Computing Systems. s. 377-378. ACM Press. ISBN 0-89791-755-3.

Nielsen, J & Landauer, T. 1993. A Mathematical Model of the Finding of Usability Problems. INTERCHI '93 Proceedings of the Conference on Human Factors in Computing Systems. s. 206-213. Addison-Wesley Longman Publishing, Boston. ISBN 0-201-58884-6.

Paronen, P. 2004a. Informaatiokooste, Tietoa Systems Gardenista sekä NAS – tuoteperheestä. Systems Garden Oy

Paronen, P. 2004b. NAS White Paper. Systems Garden Oy

Roth, E. M., Patterson, E.S. & Mumaw, R. J. 2002. Cognitive Engineering: Issues in User-Centered System Design. In J. J. Marciniak (Ed.), Encyclopedia of Software Engineering, 2nd Edition (s. 163 - 179). New York: Wiley-Interscience, John Wiley

& Sons. [Viitattu 24.10.2004] Saatavilla:
<http://emroth.home.mindspring.com/cewiley.pdf>

Rosenberg, D. 2004. The myths of usability ROI. *Interactions*, September + October 2004, Vol 11, nro 5. s. 22 – 29. ACM Press. ISSN:1072-5520/04/0900

Schranz, M. 1998. Engineering flexible World Wide Web services. *Proceedings of the 1998 ACM symposium on Applied Computing*. s. 712-718. ACM Press. ISBN 0-89791-969-6

Shapiro. C., Varian H. 1999. *Information rules: a strategic guide to network economy*. Harvard Business School Press. ISBN 0-87584-863-X

Sutcliffe, A. 2002. Assessing the Reliability of Heuristic Evaluation for Website Attractiveness and Usability. *Proceedings of the 35th Hawaii International Conference on System Sciences – 2002*. IEEE Computer Society. ISSN: 0-7695-1435

Uehling, D. 1994. *Usability Testing Handbook*. Data Systems Technology Laboratory Series DSTL-94-002. NASA, Goddard Space Flight Center, Data Systems Technology Division Software and Automation Systems Branch. [Viitattu 2.5.2005]
 Saatavilla: <http://usability.gsfc.nasa.gov/use/Usability/>

Whitefield, A., Wilson, F. & Dowell, J. 1991. A framework for human factors evaluation. *Behaviour & Information Technology*, 1991, 10(1), 65-79.

Whitefield, E., Wiggins, M. 1998. WEBDAV: IETF standard for collaborative authoring on the web. *IEEE Internet Computing*, September – October 1998, Volume 2, nro 5, s. 34-40.

Virzi, R., Sorce, J., & Herbert, L. (1993). A comparison of three usability evaluation methods: Heuristic, think-aloud, and performance testing. Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting, s. 309-313. Human Factors and Ergonomics Society.

Woolrych, A. & Cockton, G. 2001. Why and when five test users aren't enough. In Proceedings of IHM-HCI 2001 Conference Volume 2, s. 105-108. [Viitattu 6.10.2004] Saatavilla: <http://osiris.sunderland.ac.uk/~cs0gco/fiveusers.doc>

Öörni, K. 2003. What do we know about usability evaluation? A critical view. Draft send to Digilib 2003. [Viitattu 28.4.2005] Saatavilla: <http://www.student.oulu.fi/~koorni/digilib2003.pdf>

9 Liitteet

Liite A: Heuristisen arvioinnin ohjeistus

Heuristinen arviointi

Heuristinen arviointi (heuristic evaluation) on vapaamuotoinen arviointimenetelmä, jossa käytettävyyssongelmia pyritään löytämään tutkimalla käyttöliittymää ja vertaamalla sitä yleiseen suunnittelusäännöstöön eli heuristiikkaan. Tässä arvioinnissa käytetään Nielsenin ja Molichin perinteistä kymmenen kohdan heuristiikkaa.

Testin suorittaminen

Testi suoritetaan kahdessa vaiheessa. Ensimmäisessä vaiheessa jokainen käy itsenäisesti läpi ohessa esitetyt tehtävät testisivustolla osoitteessa:

<http://usability.dev.sgnet.systemsgarden.com>

käyttäjätunnus: etunimi sukunimi salasana: usability

Jokaista tehtävää tehdessä toimintoja vertaillaan suunnittelusäännöstöön. Mikäli huomaat käyttöliittymän rikkovan jotain suunnittelusäännösten kohtaa/kohtia kirjaa huomiosi erilliselle paperille tarkastelemalla ongelmaa oheisen 8 kohdan raportointimallin avulla

1. Kuvaa ongelma ja ratkaisu erillään
2. Tarjoa perustelut ongelmalle ja ratkaisulle
3. Liitä arvio ongelman vakavuudesta (kts. oheinen asteikko)
4. Pohdi avoimesti vaihtoehtoja
5. Arvioi huolellisesti kaikki ratkaisut, jotka vaativat uuden toiminnallisuuden lisäyksen sovellukseen

6. Ole tietoinen omista painotuksistasi, joita arvioijana tuot mukaan arviointiin.
7. Yritä katsoa jokaista ongelmaa usealta eri kantilta
8. Arvioi ongelmaraporttia kokonaisuutena, karsien paikallisia optimointeja ja kompromisseja yleisemmillä ratkaisuilla jotka sopivat sovellukseen kokonaisuudessaan.

Ongelmien vakavuus tulee määrittää seuraavan asteikon avulla

0. ei käytettävyysongelma
1. kosmeettinen käytettävyysongelma. Ei tarvitse korjata ellei projektissa ole ylimääräistä aikaa
2. vähäinen käytettävyysongelma. Korjaukselle tulee antaa alhainen prioriteetti
3. merkittävä käytettävyysongelma. Tärkeää korjata korkealla prioriteetilla
4. käytettävyysskatastrofi. Ehdotonta korjata ennen julkaisua

Kun kaikki ovat käyneet rauhalliseen tahtiin tehtävät läpi siirrytään toiseen vaiheeseen, jossa kaikki löydökset arvioidaan yhdessä.

Testitehtävät

- Kirjautu sisään järjestelmään
- Lisää It-osaston uutispalstalle uutinen otsikolla ”Sähköpostiohjelmien päivitys”, sisällöllä ”Vanhat Outlook 98 -ohjelmat tulee päivittää versioon 2003. Ota yhteyttä Mauri Pulkkiseen p. 3452”. Määritä myös, että uutinen poistuu automaattisesti 15.3.2005.
- Muokkaa Myynti-osaston pääsivua ja lisää sivulle teksti-osaan ”Syksyn myyntikatsaus”-teksti ja liitä sen perään liitetiedostona ”myyntikatsaus.xls”-tiedosto. (löytyy osoitteesta: \\tammi\share\Muut\Käytettävyytestin_liitetiedostot)
- Lisää Suunnittelu-osaston pääsivun alle uusi sivu, jonka navigaatiossa näkyvä otsikko on ”Viikkopalaverit”, ja varsinaisella sivulla näkyvä

otsikko on ”Suunnittelu-osaston viikkopalaverit” ja tekstinä ”Osaston viikkopalaveri pidetään neuvotteluhuoneessa Mega maanantaisin klo 10-11”. Määritä, että sivun linkki lisätään näkymään navigaatiossa ennen ”Projektit”-sivua.

- Muokkaa vielä juuri luomaasi ”Viikkopalaverit”-sivua ja lisää tekstin loppuun omalle rivilleen kuva osaston henkilöstöstä (henkilosto.jpg) (löytyy osoitteesta: \\tammi\share\Muut\Käytettävyytestin_liitetiedostot)
- Siirrä IT-osaston alisivuna oleva ”tiedotteita”-sivu kokonaisuudessaan Intranetin pääsivun alisivuksi
- Poista Myynti-osaston alisivu ”Asiakastyytyväisyys 2002”
- Lisää Intranetin pääsivulle Gallup-kysely, jonka otsikko on ”Missä haluaisit viettää talviurheilupäivän?” ja vastausvaihtoehtoina ”Peuramaan laskettelukeskuksessa, Tapiolan urheilupuistossa, Oulunkylän tekojäärädällä”. Äänestyksen tarkistusmuodon voit valita vapaasti.
- Vaihda salasanasi muotoon ”huuhaa123”
- Tutki mikä on Teppo Teikarisen käyttäjätunnus.
- Vaihda työkalupakin (valikko vasemmassa reunassa) näyttömuodoksi lista
- Kirjaudu järjestelmästä ulos

Liite B: Käytettävyydestien harjoitustehtävät

1. Kirjautu sisään järjestelmään
2. Aseta etusivun uutissyöttö keräämään uutisia kaikilta tarjollaolevilta uutispalstoilta. (Tämä tehtävä poistettiin varsinaisista käytettävyydesteistä pilottitestin arvioinnin jälkeen)
3. Lisää Myynti-osion alasivulla olevaalle keskustelupalstalle uusi keskusteluaihe otsikolla ”Myyntin tavoitteet vuonna 2005” ja tekstillä ”Tämän vuoden myyntin tavoitteet julkistettiin. Mitä mieltä olette niistä?”

Liite C: Käytettävyydestien ja heuristisen arvioinnin testitehtävät

1. Litää It-osaston uutispalstalle uutinen otsikolla ”Sähköpostiohjelmien päivitys”, sisällöllä ”Vanhat Outlook 98 -ohjelmat tulee päivittää versioon 2003. Ota yhteyttä Mauri Pulkkiseen p. 3452”. Määritä myös, että uutinen poistuu automaattisesti 15.3.2005.
2. Muokkaa Myynti-osaston pääsivua ja lisää sivulle teksti-osaan ”Syksyn myyntikatsaus”-teksti ja liitä sen perään liitetiedostona ”myyntikatsaus.xls”-tiedosto. (liitetiedosto CD:llä tai disketillä)
3. Lisää Suunnittelu-osaston pääsivun alle uusi sivu, jonka navigaatiossa näkyvä otsikko on ”Viikkopalaverit”, ja varsinaisella sivulla näkyvä otsikko on ”Suunnittelu-osaston viikkopalaverit” ja tekstinä ”Osaston viikkopalaveri pidetään neuvotteluhuoneessa Mega maanantaisin klo 10-11”. Määritä, että sivun linkki lisätään näkymään navigaatiossa ennen ”Projektit”-sivua.
4. Muokkaa vielä juuri luomaasi ”Viikkopalaverit”-sivua ja lisää tekstin loppuun omalle rivilleen kuva osaston henkilöstöstä (henkilosto.jpg) (Kuva CD:llä tai disketillä)
5. Siirrä IT-osaston alasivuna oleva ”tiedotteita”-sivu kokonaisuudessaan Intranetin pääsivun alasivuksi

6. Poista Myynti-osaston alasivu ”Asiakastyytyväisyys 2002”
7. Lisää Intranetin pääsivulle Gallup-kysely, jonka otsikko on ”Missä haluaisit viettää talviurheilupäivän?” ja vastausvaihtoehtoina ”Peuramaan laskettelukeskuksessa, Tapiolan urheilupuistossa, Oulunkylän tekojäädalla”. Äänestyksen tarkistusmuodon voit valita vapaasti.
8. Vaihda salasanasi muotoon ”huuhaa123”
9. Tutki mikä on Teppo Teikarisen käyttäjätunnus.
10. Vaihda työkalupakin (valikko vasemmassa reunassa) näyttömuodoksi lista
11. Kirjaudu järjestelmästä ulos